Limit determination: Part I

Christian Autermann

24.03.2015

General information

Contact

Christian Autermann
Lectures & Exercises
Physikzentrum 28B210
auterman@physik.rwth-aachen.de

Jan-Frederik Schulte
Exercises
Modulbau 2 202
Jan.Frederik.Schulte@rwth-aachen.de

Literature

- Volker Blobel, E. Lohrmann: „Statistische und Numerische Methoden der Datenanalyse", Teubner, ISBN-13: 978-3519032434.
- Glen Cowan: „Statistical Data Analysis", Oxford Press, ISBN-13: 978-0198501558.
- R.J. Barlow: „Statistics: A Guide to the Use of Statistical Methods in the Physical Sciences", John Wiley, ISBN-13: 978-0471922957.
- Particle Data Group: „Review of Particle Physics", Journal of Physics G: Nuclear and Particle Physics
- O. Behnke et.al. (Editor): „Data Analysis in High Energy Physics", Wiley-VCH, ISBN: 978-3527410583.
- Bjarne Stroustrup: „The C++ Programming Language", Addison-Wesley, ISBN-13: 978-0321563842.
- „The Python Tutorial", https://docs.python.org/2/tutorial/.
- Root Data Analysis Framework: „The Root User Guide 5.34", http://root.cern.ch/drupal/.
- Bjarne Stroustrup (Editor): „The C++ In Depth Series", Addison-Wesley.

Contents

Today

- Confidence levels
- Confidence level belt construction
- Confidence intervals
- Coverage

Tomorrow

- Hypotheses, decisions, and tests
- Limit calculation
- The CLs method

Overview

- Bayes theorem
  - Probability
  - Example: Bayes statistics
- Confidence Levels
- Confidence Level Belt
- Confidence Intervals
- Coverage
- Exercises

Probability

Bayes theorem

A quick reminder

Conditional probability $P(A|B)$: The probability of A to occur under the condition that B has occured.

Bayes theorem:

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

Bayes theorem

A quick reminder

Conditional probability $P(A|B)$: The probability of A to occur under the condition that B has occured.
Bayes theorem:

likelihood prior probability

posterior probability $$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$ evidence

Bayes theorem

A quick reminder

Conditional probability $P(A|B)$: The probability of A to occur under the condition that B has occured.

Bayes theorem:

likelihood prior probability

posterior probability $$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$ evidence

The following useful identity follows from the three Kolmogorov axioms:

$$P(B) = \sum_i P(B|A_i) \cdot A_i$$

for a binomial experiment this becomes:

$$= P(B|A) \cdot P(A) + P(B|!A) \cdot P(!A)$$

Bayes statistics

Example (Barlow): Subjective probability

Choose a coin from your pocket and toss it three times: It comes down head each time. The probability for this to happen is $(\frac{1}{2})^3 = \frac{1}{8}$. But could the coin be a double-headed phony (biased coin)?

Bayes statistics

Example (Barlow): Subjective probability

Choose a coin from your pocket and toss it three times: It comes down head each time. The probability for this to happen is $(\frac{1}{2})^3 = \frac{1}{8}$. But could the coin be a double-headed phony (biased coin)?

$$P(\text{phony}|\text{3 heads}) = \frac{P(\text{3 heads}|\text{phony})}{P(\text{3 heads})} \cdot P(\text{phony})$$

Bayes statistics

Example (Barlow): Subjective probability

Choose a coin from your pocket and toss it three times: It comes down head each time. The probability for this to happen is $(\frac{1}{2})^3 = \frac{1}{8}$. But could the coin be a double-headed phony (biased coin)?

$$P(\text{phony}|3\text{ heads}) \quad = \quad \frac{P(3\text{ heads}|\text{phony})}{P(3\text{ heads})} \cdot P(\text{phony})$$

Our subjective prior: One of a million randomly choosen coins is a phony.

And $P(3\text{ heads}) = P(3\text{ heads}|\text{fair}) \cdot P(\text{fair}) + P(3\text{ heads}|\text{phony}) \cdot P(\text{phony})$.

Bayes statistics

Example (Barlow): Subjective probability

Choose a coin from your pocket and toss it three times: It comes down head each time. The probability for this to happen is $(\frac{1}{2})^3 = \frac{1}{8}$. But could the coin be a double-headed phony (biased coin)?

$$
\begin{aligned}
P(\text{phony}|\text{3 heads}) &= \frac{P(\text{3 heads}|\text{phony})}{P(\text{3 heads})} \cdot P(\text{phony}) \\
&= \frac{1 \cdot 10^{-6}}{0.125 \cdot (1 - 10^{-6}) + 1 \cdot 10^{-6}} \\
&= 8 \cdot 10^{-6}
\end{aligned}
$$

Our subjective prior: One of a million randomly choosen coins is a phony.
And $P(\text{3 heads}) = P(\text{3 heads}|\text{fair}) \cdot P(\text{fair}) + P(\text{3 heads}|\text{phony}) \cdot P(\text{phony})$.

Bayes statistics

Honest Harry

„What about this little baby? Seventeen previous owners, twice round the clock, drinks like a fish, goes like a tortoise...”

http://www.cartoonstock.com

Example (Barlow): Subjective probability

Now you are on the Reeperbahn with Honest Harry, the used car salesman, who suggests to toss a coin to see who pays for the drinks.

Again, it comes down head three times.

Now the prior probability that this coin is a phony is not 10^{-6} but larger, lets say 5%.
Is the coin a phony?

Bayes statistics

Honest Harry

„What about this little baby? Seventeen previous owners, twice round the clock, drinks like a fish, goes like a tortoise..."

http://www.cartoonstock.com

Example (Barlow): Subjective probability

Now you are on the Reeperbahn with Honest Harry, the used car salesman, who suggests to toss a coin to see who pays for the drinks.

Again, it comes down head three times.

Now the prior probability that this coin is a phony is not 10^{-6} but larger, lets say 5%.
Is the coin a phony?

$$P(\text{phony}|3\text{ heads}) = \frac{1 \cdot 0.05}{0.125 \cdot (0.95) + 1 \cdot 0.05} = 30\%$$

Overview

Confidence levels

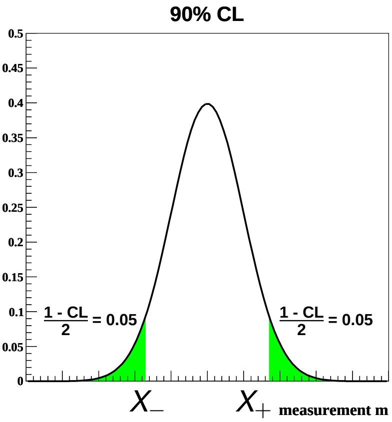
http://buildingabrandonline.com/Beinspired/thrive-with-unstoppable-confidence-part-3/

Confidence level intervals

Given a precisely known true value μ of a certain property (e.g. the weight of cereal packets), we can ask:

- What is the weight-range into which a certain amount (e.g. 90%) of measurements x_i will fall?

90% central confidence interval for a Gaussian



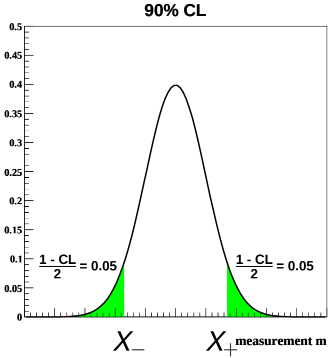
The value of the measurement m lies in the interval $X_- \ldots X_+$ in „CL"% of the time.

$\Longleftrightarrow$ The statement „m will lie in the interval $X_- \ldots X_+$" has CL confidence.

Central confidence intervals for a Gaussian distributions:

$$P(X_- \geq x \geq X_+) = \int_{X_-}^{X_+} P(x)dx = CL$$

x: measurement, $X_\pm$ limits of the confidence interval.



Common values:

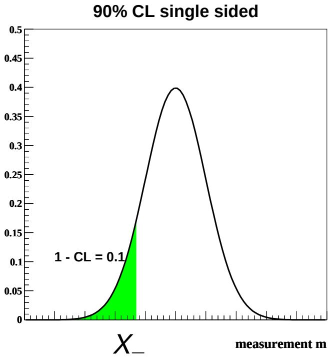
1σ	$\hat{=}$	68.27%	1.6449σ	$\hat{=}$	90%
2σ	$\hat{=}$	95.45%	1.9600σ	$\hat{=}$	95%
3σ	$\hat{=}$	99.73%	2.5758σ	$\hat{=}$	99%
5σ	$\hat{=}$	99.99994%			

Single sided confidence intervals for a Gaussian distributions:

$$P(x \geq X_+) = \int_{-\infty}^{X_+} P(x)dx = CL_{\text{upper}}$$

$$P(X_- \geq x) = \int_{X_-}^{\infty} P(x)dx = CL_{\text{lower}}$$

x: measurement, $X_\pm$ single sided limits of the confidence interval.



Common values:

1σ	$\hat{=}$	84.13%	1.2816σ	$\hat{=}$	90%
2σ	$\hat{=}$	97.72%	1.6449σ	$\hat{=}$	95%
3σ	$\hat{=}$	99.87%	2.3263σ	$\hat{=}$	99%
5σ	$\hat{=}$	99.99997%			

Example single sided confidence interval: Journey to work

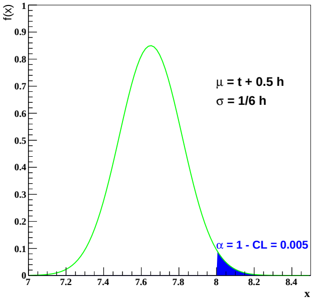
An employee needs to be at work at 8:00 o'clock sharp. The journey takes 30 minutes on average, with a Gaussian uncertainty of $\sigma = 10$ minutes due to varying traffic.

When must he leave home to be late only once a year ($\sim 0.5\%$)?

Example single sided confidence interval: Journey to work

An employee needs to be at work at 8:00 o'clock sharp. The journey takes 30 minutes on average, with a Gaussian uncertainty of $\sigma = 10$ minutes due to varying traffic.

When must he leave home to be late only once a year ($\sim 0.5\%$)?



$$\int_{X_{up}}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = 1 - CL$$

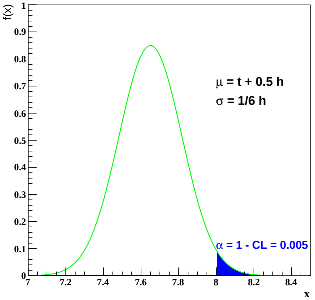
$$\begin{array}{lcl} \text{mean} \quad \mu & = & t + 30 \text{ minutes} \\ \text{limit} \quad X_{up} & = & 8:00 \text{ o'clock} \\ \text{width} \quad \sigma & = & 10 \text{ minutes} \end{array}$$

Example single sided confidence interval: Journey to work

An employee needs to be at work at 8:00 o'clock sharp. The journey takes 30 minutes on average, with a Gaussian uncertainty of $\sigma = 10$ minutes due to varying traffic.

When must he leave home to be late only once a year ($\sim 0.5\%$)?

Single sided limit: ($2.3\sigma \hat{=} 99.0\%$, $3\sigma \hat{=} 99.87\%$) $\longrightarrow 99.5\% \approx 2.5\sigma$



$$\int_{X_{up}}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = 1 - CL$$

$$\begin{array}{llcl} \text{mean} & \mu & = & t + 30 \text{ minutes} \\ \text{limit} & X_{up} & = & 8:00 \text{ o'clock} \\ \text{width} & \sigma & = & 10 \text{ minutes} \end{array}$$

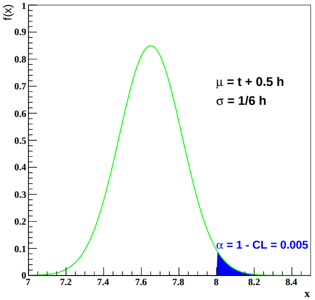
Example single sided confidence interval: Journey to work

An employee needs to be at work at 8:00 o'clock sharp. The journey takes 30 minutes on average, with a Gaussian uncertainty of $\sigma = 10$ minutes due to varying traffic.

When must he leave home to be late only once a year ($\sim 0.5\%$)?

Single sided limit: ($2.3\sigma \hat{=} 99.0\%$, $3\sigma \hat{=} 99.87\%$) $\longrightarrow$ $99.5\% \approx 2.5\sigma$

He has to leave at t $\approx$ 8:00 - 0:30 - 0:10·2.5 =7:05 o'clock!



$$\int_{X_{up}}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = 1 - CL$$

$$\begin{array}{llcl} \text{mean} & \mu & = & t + 30 \text{ minutes} \\ \text{limit} & X_{up} & = & 8:00 \text{ o'clock} \\ \text{width} & \sigma & = & 10 \text{ minutes} \end{array}$$

Confidence level intervals

With a true value μ of a certain property (e.g. the weight of cereal packets) with a width σ, we can ask:

- Given one measurement x, what could we say about the true value μ?

Simply turning it around, to say that μ lies in the interval $x - \sigma \ldots x + \sigma$ is naive, because it contains hidden assumptions.

$\Longrightarrow$ Confidence belt to translate a measurement into a confidence interval.

Overview

- Bayes theorem
- Confidence Levels
- Confidence Level Belt
  - Frequentistic confidence level belt construction
- Confidence Intervals
- Coverage
- Exercises

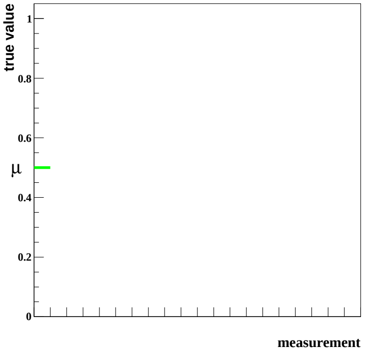
Confidence Belt

http://www.inewidea.com ($90)

Confidence Level Belt

Neyman construction for Confidence Level intervals

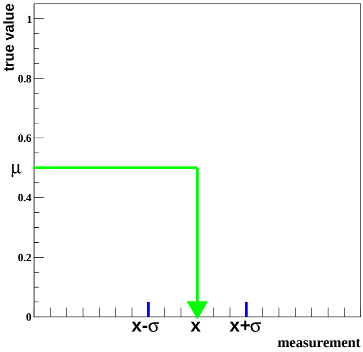
Given a particular true value μ



Confidence Level Belt

Neyman construction for Confidence Level intervals

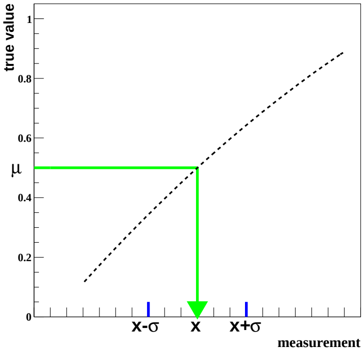
Given a particular true value μ , there is a probability density function $P(\mu, \sigma)$ that defines the most probable measurement x, and the interval $x - \sigma \ldots x + \sigma$ into which the measurements will fall with a given CL.



Confidence Level Belt

Neyman construction for Confidence Level intervals

Given a particular true value μ , there is a probability density function $P(\mu, \sigma)$ that defines the most probable measurement x, and the interval $x - \sigma \ldots x + \sigma$ into which the measurements will fall with a given CL.

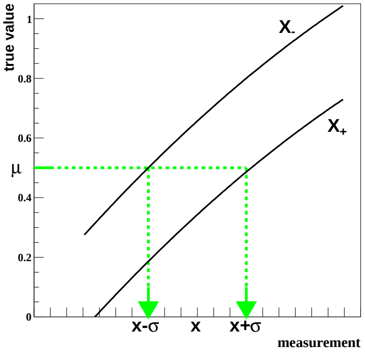


1. For a different μ there are different measurements x and limits $x \pm \sigma$.

Confidence Level Belt

Neyman construction for Confidence Level intervals

Given a particular true value μ , there is a probability density function $P(\mu, \sigma)$ that defines the most probable measurement x, and the interval $x - \sigma \ldots x + \sigma$ into which the measurements will fall with a given CL.

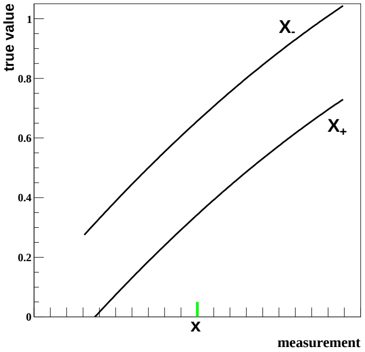


1. For a different μ there are different measurements x and limits $x \pm \sigma$.
2. The measurement limits $x - \sigma$ and $x + \sigma$ can be considered as functions from the true value μ.

Confidence Level Belt

Neyman construction for Confidence Level intervals

Given a particular true value μ , there is a probability density function $P(\mu, \sigma)$ that defines the most probable measurement x, and the interval $x - \sigma \ldots x + \sigma$ into which the measurements will fall with a given CL.

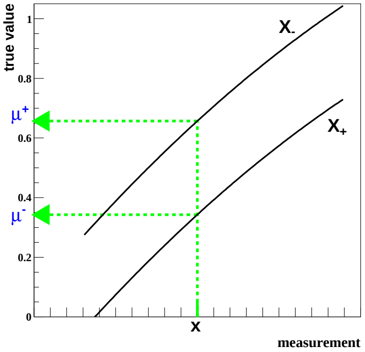


1. For a different μ there are different measurements x and limits $x \pm \sigma$.
2. The measurement limits $x - \sigma$ and $x + \sigma$ can be considered as functions from the true value μ.
3. The functions $X_-(\mu)$ and $X_+(\mu)$ are the **confidence belt.**

Confidence Level Belt

Neyman construction for Confidence Level intervals

Given a particular true value μ , there is a probability density function $P(\mu, \sigma)$ that defines the most probable measurement x, and the interval $x - \sigma \ldots x + \sigma$ into which the measurements will fall with a given CL.

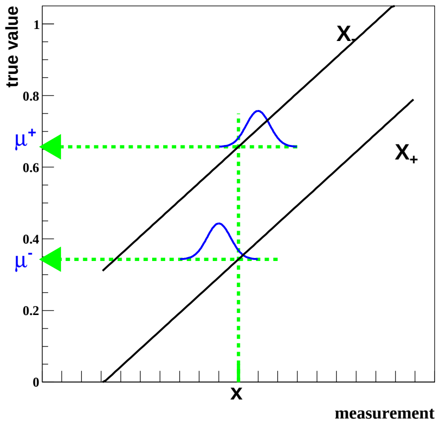


Given a measurement x a confidence interval for the **true value** $\mu^- \ldots \mu^+$ can be constructed from the confidence belt.

The confidence belt is *constructed horizontally* using the known probability density for all possible true values μ. Having a measurement x, it is *read vertically.*

The $\mu^- \ldots \mu^+$ enclose with CL probability the true value μ.

Example: Confidence Belt for a Gaussian



For **Gaussian distributions** the conversion from the horizontal measurement confidence interval $x^- \ldots x^+$ to the vertical true confidence interval $\mu^- \ldots \mu^+$ is simple: The confidence belt X_-, X_+ becomes two straight lines with unit gradient.

$x_\pm = \mu \pm n \cdot \sigma$	when constructed horizontally
$\mu_\pm = x \pm n \cdot \sigma$	when read vertically

With $n = 1$ for $CL = 68\%$, etc...

Overview

- Bayes theorem
- Confidence Levels
- Confidence Level Belt
- Confidence Intervals
  - Binomial Confidence Intervals
  - Poisson Confidence Intervals
  - Constrained Confidence Intervals
- Coverage
- Exercises

Binomial Confidence Intervals

„Coin-flip" experiments

Binomial experiments have only two possible outcomes. While the true value μ is continous the observed value is discrete. The confidence integrals become summations.
For m successes in n binomial tries, the limits p_- and p_+ of the confidence interval are found by:

$$\sum_{r=0}^{m-1} B(\mu, p_-, n) \leq \frac{\mathrm{CL}}{2} \qquad \sum_{r=m+1}^{n} B(\mu, p_+, n) \leq \frac{\mathrm{CL}}{2}$$

where

$$B(\mu, p, n) = \left(\begin{array}{c} n \\ k \end{array} \right) p^{\mu} (1-p)^{n-\mu}$$

is the binomial distribution.

Poisson Confidence Intervals

„Coin-flip" experiments

A Poisson-distribution is a approximation for a binomial for large n and small probabilities p, i.e. $n \to \infty$ and $p \to 0$.

$$P(k, \lambda) = \frac{\lambda^k}{k!} e^{-\lambda}$$

where k is the number of successes per interval, and λ the true expectation.
The limits of the confidence interval become:

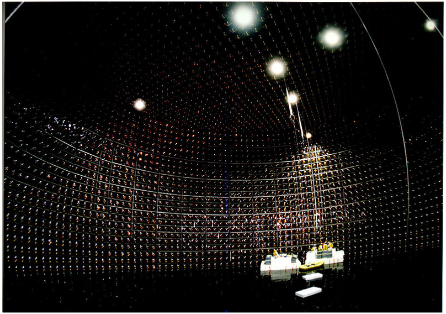
$$\sum_{r=0}^{k-1} P(r, \lambda_-) \leq \frac{\text{CL}}{2} \qquad \sum_{r=n+1}^{\infty} P(r, \lambda_+) \leq \frac{\text{CL}}{2}$$

(lower) (upper)

Poisson Confidence Intervals

Example: Proton decay

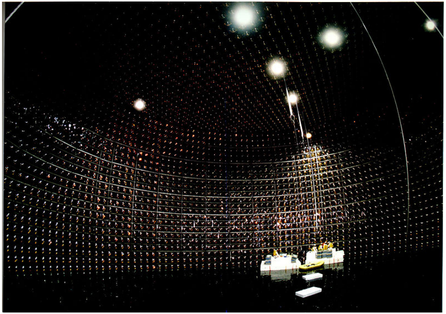
In Super-Kamiokande with 50 000 tons of water, less than s proton-decay candidate events per year are observed. What is the 95% CL interval for proton-decays and the proton half-life, *assuming no background* events and $s = 1$ found event per year?



Poisson Confidence Intervals

Example: Proton decay

In Super-Kamiokande with 50 000 tons of water, less than s proton-decay candidate events per year are observed. What is the 95% CL interval for proton-decays and the proton half-life, *assuming no background* events and $s = 1$ found event per year?



- Number of protons in 50 ktons of water: $N = 1.65 \cdot 10^{34}$

Poisson Confidence Intervals

Example: Proton decay

In Super-Kamiokande with 50 000 tons of water, less than s proton-decay candidate events per year are observed. What is the 95% CL interval for proton-decays and the proton half-life, *assuming no background* events and $s = 1$ found event per year?

Poisson limits:

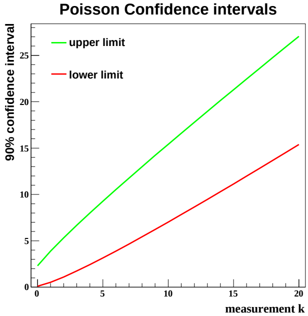
	Lower			Upper		
n	90%	95%	99%	90%	95%	99%
0	–	–	–	2.30	3.00	4.61
1	0.11	0.05	0.01	3.89	4.74	6.64
2	0.53	0.36	0.15	5.32	6.30	8.41
3	1.10	0.82	0.44	6.68	7.75	10.05
4	1.74	1.37	0.82	7.99	9.15	11.60
5	2.43	1.97	1.28	9.27	10.51	13.11
6	3.15	2.61	1.79	10.53	11.84	14.57
7	3.89	3.29	2.33	11.77	13.15	16.00
8	4.66	3.98	2.91	12.99	14.43	17.40
9	5.43	4.70	3.51	14.21	15.71	18.78
10	6.22	5.43	4.13	15.41	16.96	20.14

- Number of protons in 50 ktons of water: $N = 1.65 \cdot 10^{34}$
- 95% CL interval, e.g. for 1 event: $\mathrm{CL}_{dn} = 0.05$, $\mathrm{CL}_{up} = 4.74$.

Poisson Confidence Intervals

Example: Proton decay

In Super-Kamiokande with 50 000 tons of water, less than s proton-decay candidate events per year are observed. What is the 95% CL interval for proton-decays and the proton half-life, *assuming no background* events and $s = 1$ found event per year?



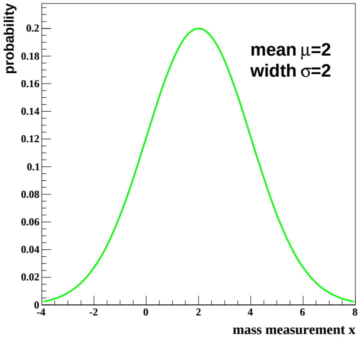
- Number of protons in 50 ktons of water: $N = 1.65 \cdot 10^{34}$
- 95% CL interval, e.g. for 1 event: $\mathrm{CL}_{dn} = 0.05$, $\mathrm{CL}_{up} = 4.74$.
- Prob. one decay/year: $P = \frac{CL_{dn}}{N} = 3.03 \cdot 10^{-36} \ldots 2.87 \cdot 10^{-34}$ and mean lifetime interval: $3.48 \cdot 10^{33} < \tau = \frac{1}{P} < 3.3 \cdot 10^{35}$ years.

Constrained Confidence Intervals

Constrained Gaussian Distributions

Given a measurement x with resolution σ we want to find the limits of the confidence intervals of the true underlying variable μ, which we know must be within a specific interval.

Example: Measuring a mass x, which we know must be positive.

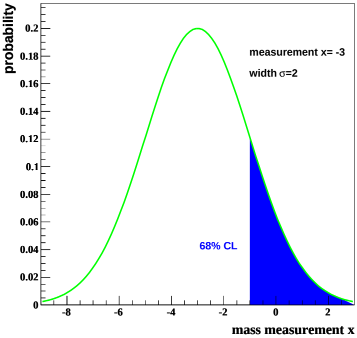


Constrained Confidence Intervals

Constrained Gaussian Distributions

Given a measurement x with resolution σ we want to find the limits of the confidence intervals of the true underlying variable μ, which we know must be within a specific interval.

Example: Measuring a mass x, which we know must be positive.
Some measurements lead to a negative upper mass limit, which is absurd.



Constrained Confidence Intervals

Constrained Gaussian Distributions

Bayes statistics allows to incorporate our prior knowledge about the true value μ.

Example: Mass measurement, μ constrained to positive values

Constrained Confidence Intervals

Constrained Gaussian Distributions

Bayes statistics allows to incorporate our prior knowledge about the true value μ.

Example: Mass measurement, μ constrained to positive values
Prior: $P(\mu) = 1$ if $\mu \geq 0$, $P(\mu) = 0$ else.

$$P(\mu_{up}|x) \quad = \quad \frac{P(x|\mu)}{P(x)} \cdot P(\mu)$$

Constrained Confidence Intervals

Constrained Gaussian Distributions

Bayes statistics allows to incorporate our prior knowledge about the true value μ.

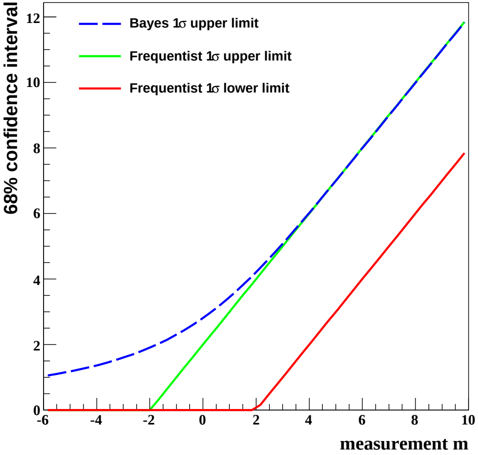
Example: Mass measurement, μ constrained to positive values
Prior: $P(\mu) = 1$ if $\mu \geq 0$, $P(\mu) = 0$ else.

$$
\begin{aligned}
P(\mu_{up}|x) &= \frac{P(x|\mu)}{P(x)} \cdot P(\mu) \\
&= \frac{\int_{-\infty}^{\mu_{up}} \mathrm{Gauss}(\sigma, x - x')dx'}{\int_0^{\infty} \mathrm{Gauss}(\sigma, x - x')dx'} \times \begin{cases} 1 & \text{for} \quad x > 0 \\ 0 & \text{else} \end{cases} \quad = 1 - \frac{CL}{2}
\end{aligned}
$$

$\Rightarrow \mu_{\text{up}}$ @ CL confidence level, (μ_{lower} equivalently) by solving the above equation.

Constrained Confidence Intervals

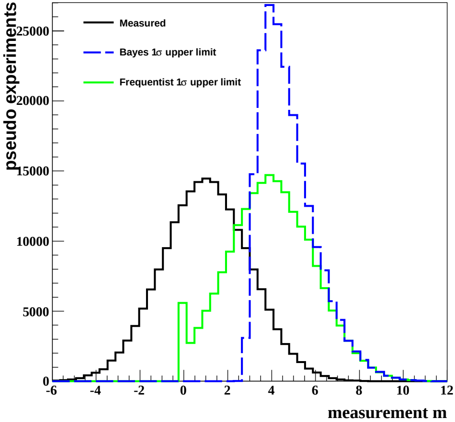
Frequentists and Bayes confidence belt



Shown is the 68% confidence belt for Bayes using a flat prior for $P(\mu)$ (constraint to positive values), and a „normal” Frequentists approach.

Constrained Confidence Intervals

Upper limits for many pseudo-experiments



For a large number of pseudo experiments the measured mass x, and the upper limits obtained from Bayes with flat prior and the Frequentists approach are shown in the plot.

While the Frequentists upper limit is a shifted Gauss truncated at 0, the Bayes upper limit is here constrained to meaningful values.

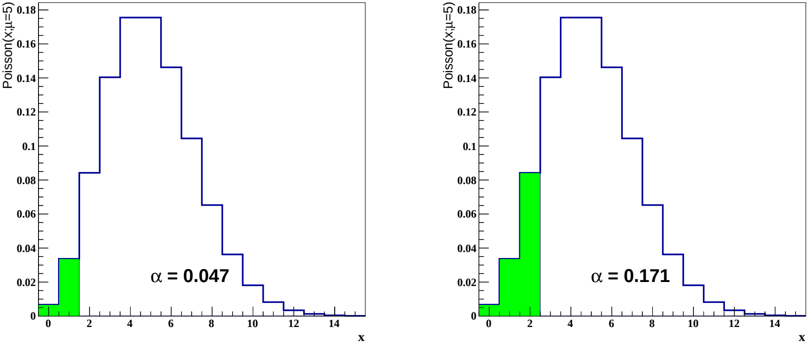
Overview

- Bayes theorem
- Confidence Levels
- Confidence Level Belt
- Confidence Intervals
- Coverage
  - Example: Poisson
- Exercises

Coverage: Poisson

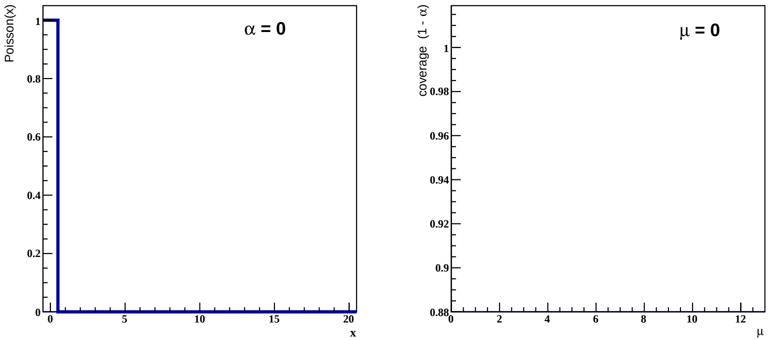
Poisson: discrete distribution

The **coverage probability** of a confidence interval is the proportion of the time that the interval contains the true value of interest



Coverage

- Single-sided 90% C.L. limit on Poisson mean



PDF reader with Java (e.g. Adobe Acrobat) necessary for animation

Conclusions Confidence Intervals

Recap Confidence Intervals

- Frequentistic limits can have unphysical values, though they are strictly correct
- 95% CL limits are not true 1 out of 20 times, by definition
- The coverage of a frequentistic limit might differ from the stated confidence level. Frequentistic limit can be conservative.
- Baysian limits can avoid these problems: Coverage is correct and the limits can be constraint to physical meaningful values
- Feldman-Cousins suggested a method to fix Frequentistic limits (Phys. Rev. D 57 (1998) 3873)

Overview

- Bayes theorem
- Confidence Levels
- Confidence Level Belt
- Confidence Intervals
- Coverage
- Exercises
  - Task 1: Thermometers
  - Task 2: Upper limit
  - Optional: Leukemia

Task 1: Thermometers

Task 1: Thermometers

A company produces clinical thermometers

- From testing a sample of thermometers it is observed that the results from different thermometers spread approximately according to a normal distribution with a sigma of 0.1 degree celsius. Estimate how many of 10000 produced thermometers will show a temperature which is:
  - more than 0.3 degrees wrong (Note: can be either too low or two high)?
  - more than +0.3 degrees wrong?
  - more than 0.4 degrees wrong?
  - more than +0.4 degrees wrong?
- If less than 5% of the thermometers should be wrong by more than 0.1 degree - than to which precision (sigma) should the thermometers be calibrated?

Solution to task 1: Thermometers

Solution to task 1: Thermometers

- more than 0.3 degrees: corresponds to 3σ. The Two-sided CL for 3σ is $2.7 \cdot 10^{-3}$. Therefore, 27 out of 10000 thermometers are expected to deviate this much or more.

Solution to task 1: Thermometers

Solution to task 1: Thermometers

- more than 0.3 degrees: corresponds to 3σ. The Two-sided CL for 3σ is $2.7 \cdot 10^{-3}$. Therefore, 27 out of 10000 thermometers are expected to deviate this much or more.
- more than +0.3 degrees: One-sided confidence interval, which is half as large, therefore 13.5 thermometers are expected to deviate by +0.3 degrees

Solution to task 1: Thermometers

Solution to task 1: Thermometers

- more than 0.3 degrees: corresponds to 3σ. The Two-sided CL for 3σ is $2.7 \cdot 10^{-3}$. Therefore, 27 out of 10000 thermometers are expected to deviate this much or more.
- more than +0.3 degrees: One-sided confidence interval, which is half as large, therefore 13.5 thermometers are expected to deviate by +0.3 degrees
- more than 0.4 degrees: The CL is $6.3 \cdot 10^{-6}$, corresponding to 0.63 thermometers.

Solution to task 1: Thermometers

Solution to task 1: Thermometers

- more than 0.3 degrees: corresponds to 3σ. The Two-sided CL for 3σ is $2.7 \cdot 10^{-3}$. Therefore, 27 out of 10000 thermometers are expected to deviate this much or more.
- more than +0.3 degrees: One-sided confidence interval, which is half as large, therefore 13.5 thermometers are expected to deviate by +0.3 degrees
- more than 0.4 degrees: The CL is $6.3 \cdot 10^{-6}$, corresponding to 0.63 thermometers.
- more than +0.4 degrees: One-sided CLs, i.e. 0.32 thermometers

Solution to task 1: Thermometers

Solution to task 1: Thermometers

- 5% corresponds to (approximately) 2σ, therefore

$$0.1 \text{ degree} \,\hat{=}\, 2\sigma$$

The produced thermometers should therefore follow a nomal distribution with width $\sigma \approx 0.05$ degrees.

Task 2: Upper limit

limit for signal + small background (Frequentist approach)

Most general, the data consists of signal and background such that $\mu = \mu_{sig} + \mu_{bkg}$. Here μ_{sig} and μ_{bkg} are the Poisson parameters for signal and background, respectively. Determine 90% CL upper limits of μ_{sig} for the following cases with a given N_{obs} and known μ_{bkg}

- $\mu_{bkg} = 0$, $N_{obs} = 2$
- $\mu_{bkg} = 1$, $N_{obs} = 2$
- $\mu_{bkg} = 3$, $N_{obs} = 0$

Hint: The relevant formula is $p(\mu, N_{obs}) = \sum_{i \le N_{obs}} e^{-\mu} \frac{\mu^i}{i!} = 10\%$, where μ has to be replaced with $\mu_{sig} + \mu_{bkg}$. See the figure for of the confidence intervals for different μ_{sig} and N_{obs} to solve subexercise a)
Note: $p(\mu, N_{obs} = 0) = e^{-\mu}$.

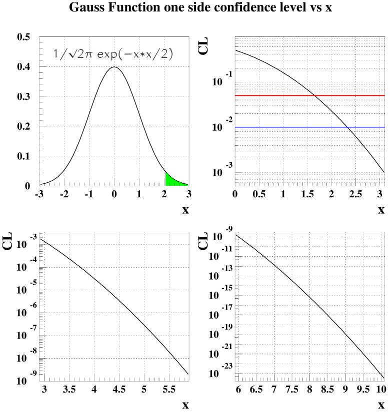
Solution to task 2: Upper limit

Solution to task 2: Upper limit

1. Consider the bin with Nobs = 2. The values are
   $p(\mu = 5, 2) = 0.12$ and
   $p(\mu = 6, 2) = 0.06$.
   The 90% CL limit is therefore at $\mu_{sig} \approx 5.3$.
2. Now μ_{bkg} is 1, therefore $\mu_{sig} = 5.3 - 1 = 4.3$.
3. The probability is $p = e^{-(\mu_{sig}+3)} = 0.1$. The signal strength μ_{sig} ougth to be negative, therefore $\mu_{sig} = 0$.

Appendix 1 - One sided Gaussian confidence levels

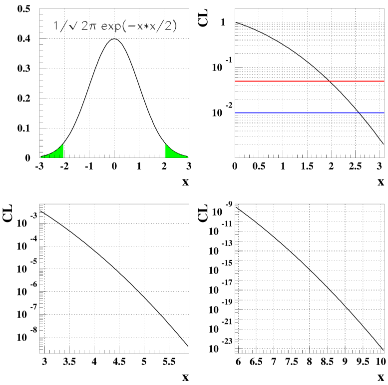
$$\mathrm{CL}(x) = \int_x^\infty dx' \, \frac{1}{\sqrt{2\pi}} e^{-x'^2/2}$$



Appendix 2 - Two sided Gaussian confidence levels

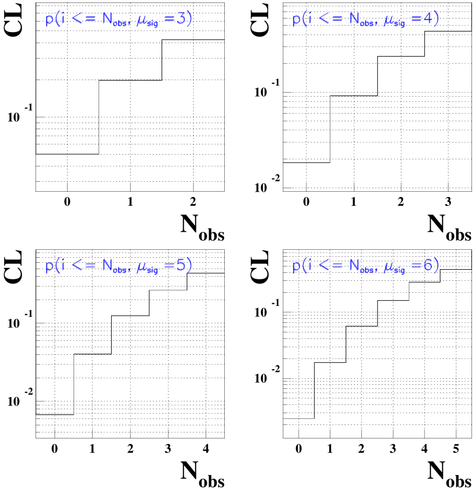
$$\mathrm{CL}(x) = 2 \int_{x}^{\infty} dx' \, \frac{1}{\sqrt{2\pi}} e^{-x'^2/2}$$

Gauss Function two side confidence level vs x



Appendix 3 - Poisson confidence levels

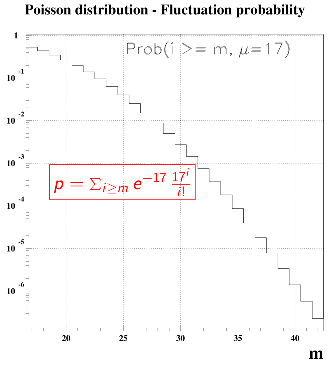
Poisson distr. - Downward fluctuation probability



Optional: Leukemia cases close to nuclear power plants

Researchers from Mainz (Maria Blettner et al) observed that in a 5 km surrounding of nuclear power plants 37 children contracted leukemia (in the years 1980-2003), while the statistical average in the population is 17. Determine the probability for a statistical fluctuation from 17 to $\geq$37.

- Use the exact poisson probabilities shown in the figure
- Approximate the distribution by a Gaussian with $\mu = 17$ and $\sigma = \sqrt{17}$. Use the CL curves for a Gaussian to determine the fluctuation probability.



Optional: Solution Leukemia

Leukemia cases close to nuclear power plants - solution

- Simply reading off the figure: $p = 2 \cdot 10^{-5}$
- Deviation in number of σ: (37-17)/$\sqrt{17}$ = 4.85 $\rightarrow$ CL = $6 \cdot 10^{-7}$

The difference between both estimates is due to the fact that the Poisson distribution has more tails towards larger numbers compared to the Gaussian. However, in both cases, the fluctuation probability is very low such than one can conclude there is a significant increase in the cancer risk close to nuclear power plants.

Further information:

- The results are vehemently disputed in other publications