# Limits in High Energy Physics

Olaf Behnke, DESY

Lecture/tutorial for the
INFN Statistics School

Ischia 9 May, 2017

Based in parts on the lecture by Stefan Schmitt, DESY,
@Terascale Statistics School, Mar 2013

# Outline

- 1) Basic interval estimation (Bayesian, Frequentist) for gaussian
- 2) Poisson without & with background, expected limit, $CL_s$
- 3) L-ratio based signal limits: Feldman-Cousins, profile likelihood
- Lecture is interleaved by exercises. Discuss solutions in the lecture
- ROOT macros for exercises:

  `www.desy.de/~sschmitt/LimitStatSchool2013/macros`

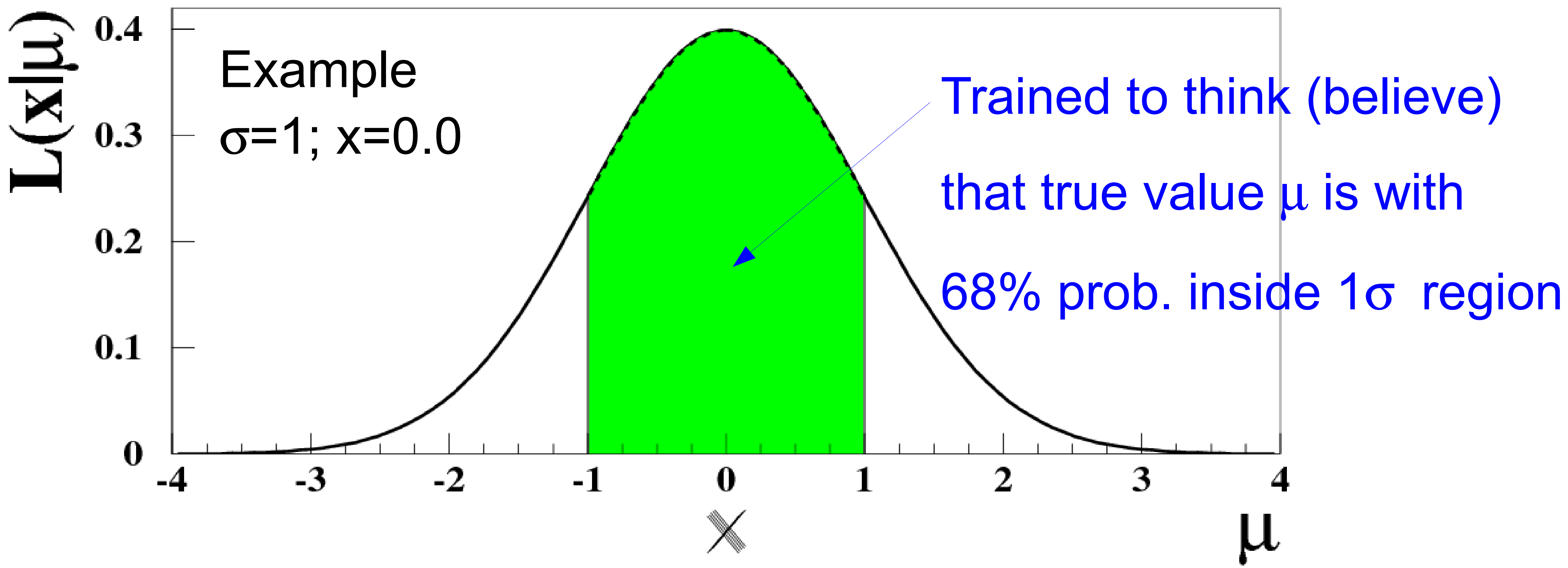
- If available, use wget:

```
wget -N -nd www.desy.de/~sschmitt/LimitStatSchool2013/macros.list
wget -N -nd -i macros.list
```

# 1) Basic interval estimation (Bayesian, Frequentist) for gaussian

# Aren't we all Bayesians?

- Measurement with gaussian uncertainty:
  - μ = true value; x = measured value; $L(x|\mu) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$



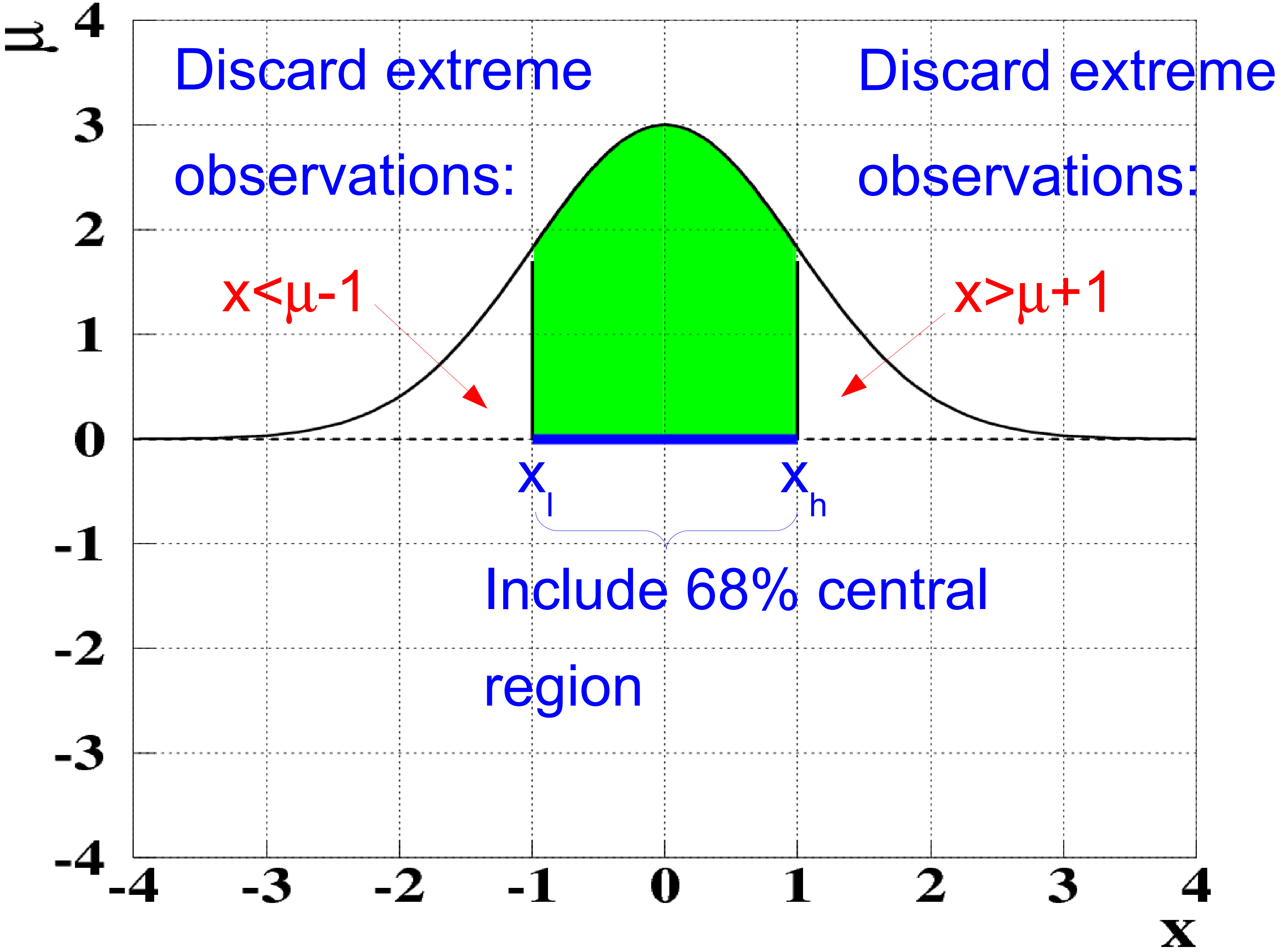

- Bayesian Posterior density: $p(\mu|x) = \frac{L(x|\mu)\pi(\mu)}{\underbrace{\int L(x|\mu)\pi(\mu)d\mu}_{=\pi(x)}}$

# Frequentist 68% C.L. central interval

- Task: Chose interval $[\mu_l,\mu_h]_x$ dependent on x such that in 68% of repeated experiments true value $\mu$ is included in the intervals
- "easy to do": define for assumed true $\mu$ interval $[x_l,x_h]_\mu$ with
  - $P(x \in [x_l,x_h]_\mu) = 68\%$
  - If $x \in [x_l,x_h]_\mu$ declare $\mu$ to be part of $[\mu_l,\mu_h]_x$ otherwise out
- $[x_l,x_h]_\mu$ construction vs true $\mu \rightarrow$ Neyman Belt (or band)

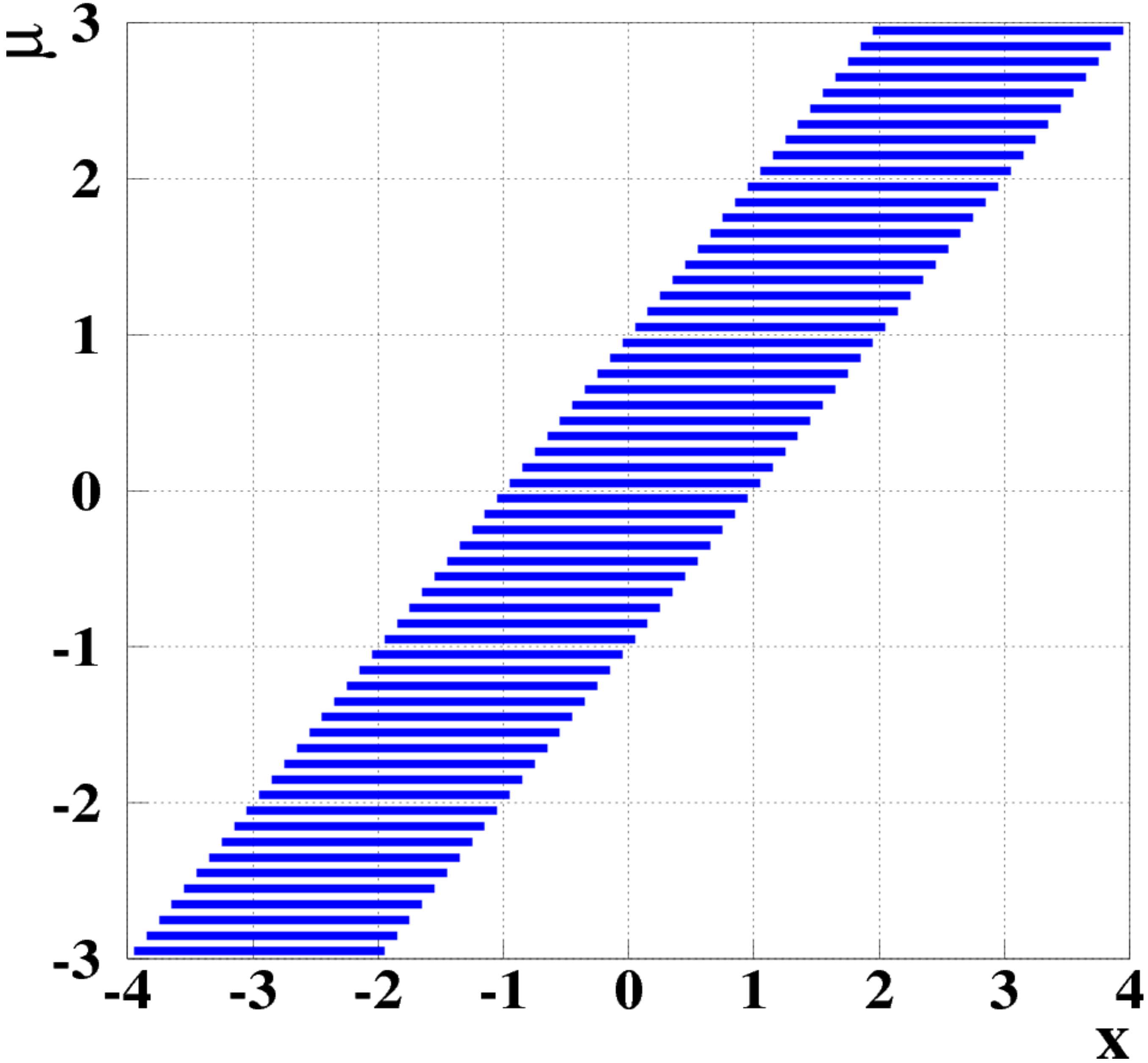
# Neyman Band for Gaussian (central interval)

- Step 1: Choose $[\, x_l, x_h\,]_\mu$ for one $\mu$ value (example: $\mu$=0)



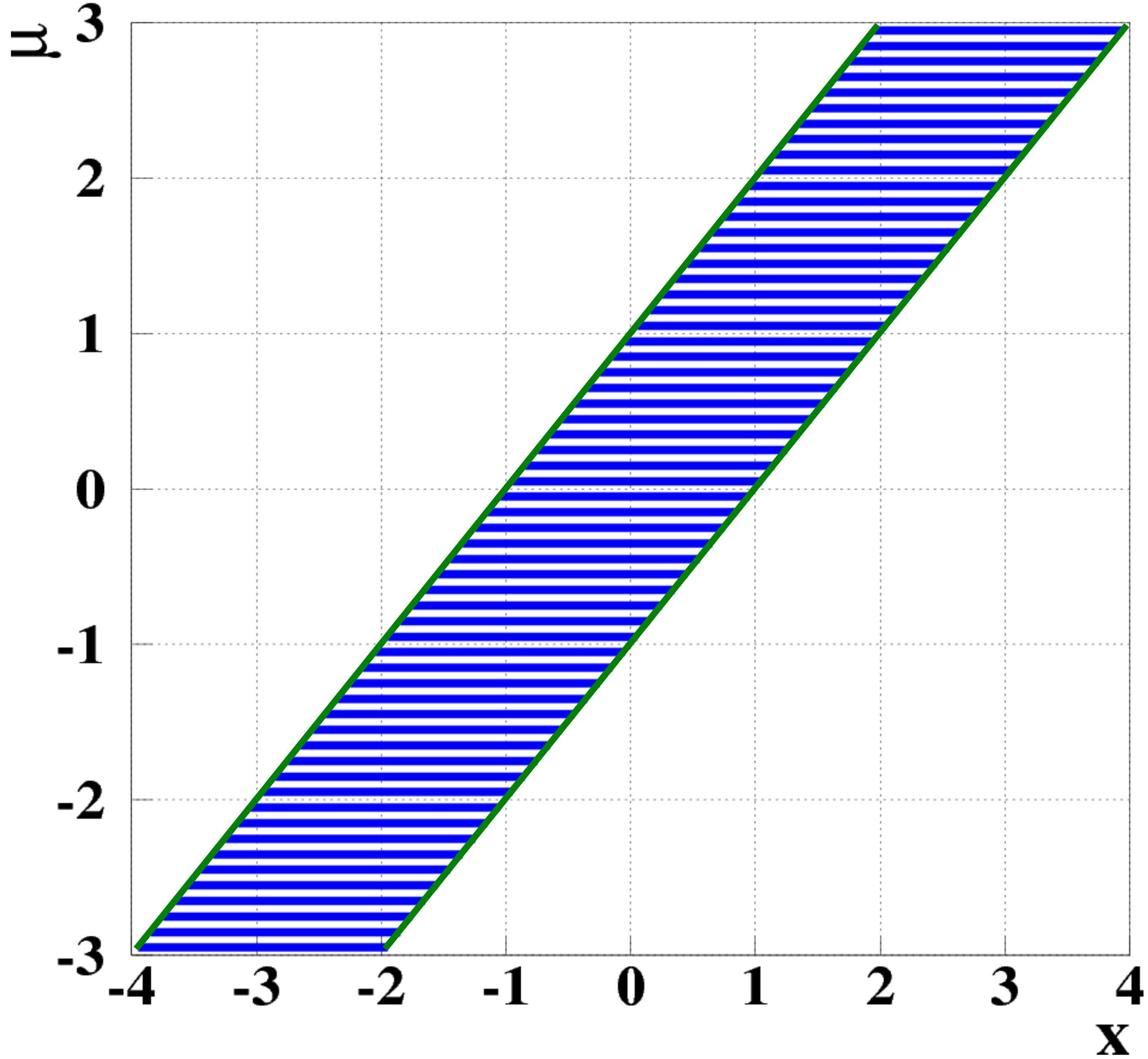
# Neyman Band for Gaussian (central interval)

- Step 2: plot $[\, x_l, x_h \,]_\mu$ vs $\mu$



# Neyman Band for Gaussian (central interval)

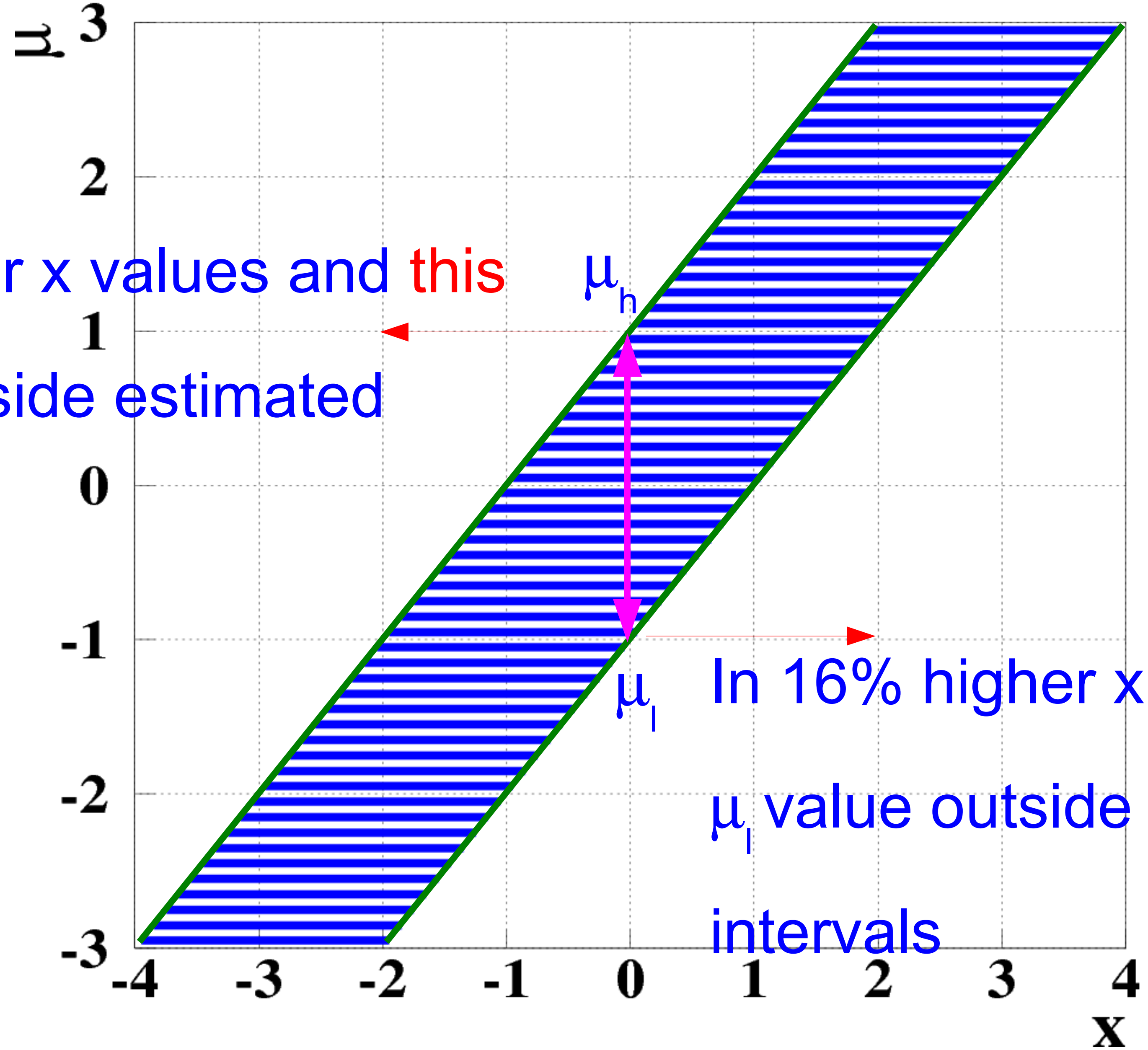
- Step 3: plot $[\, x_l, x_h \,]_\mu$ vs $\mu$; connect edges



# Neyman Band for Gaussian (central interval)

- Step 4: plot $[x_l, x_h]_\mu$ vs $\mu$; for given x: max & min $\mu \rightarrow [\mu_l, \mu_h]_x$

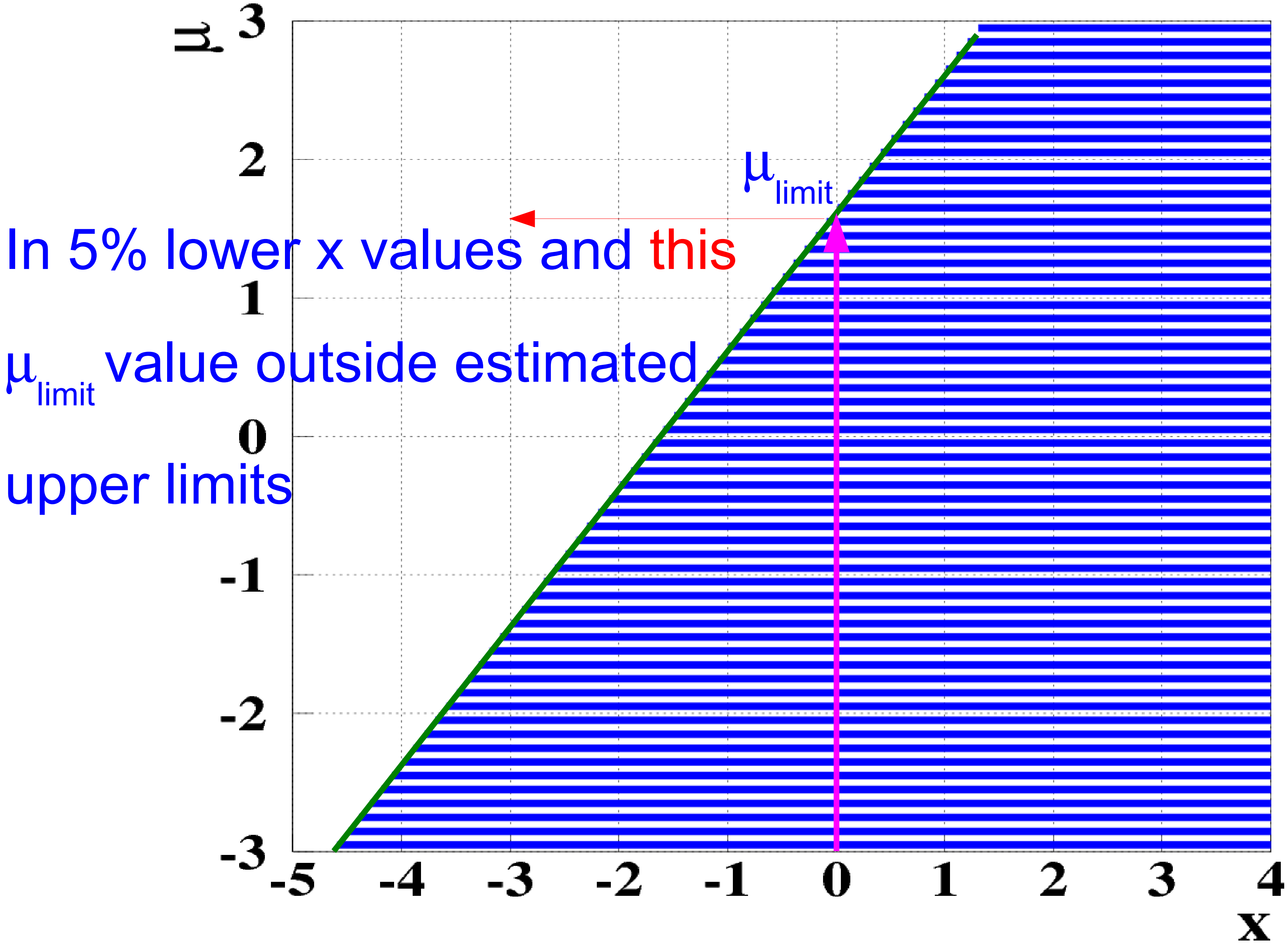
In 16% lower x values and this $\mu_h$ value outside estimated intervals




In 16% higher x values and this $\mu_l$ value outside estimated intervals

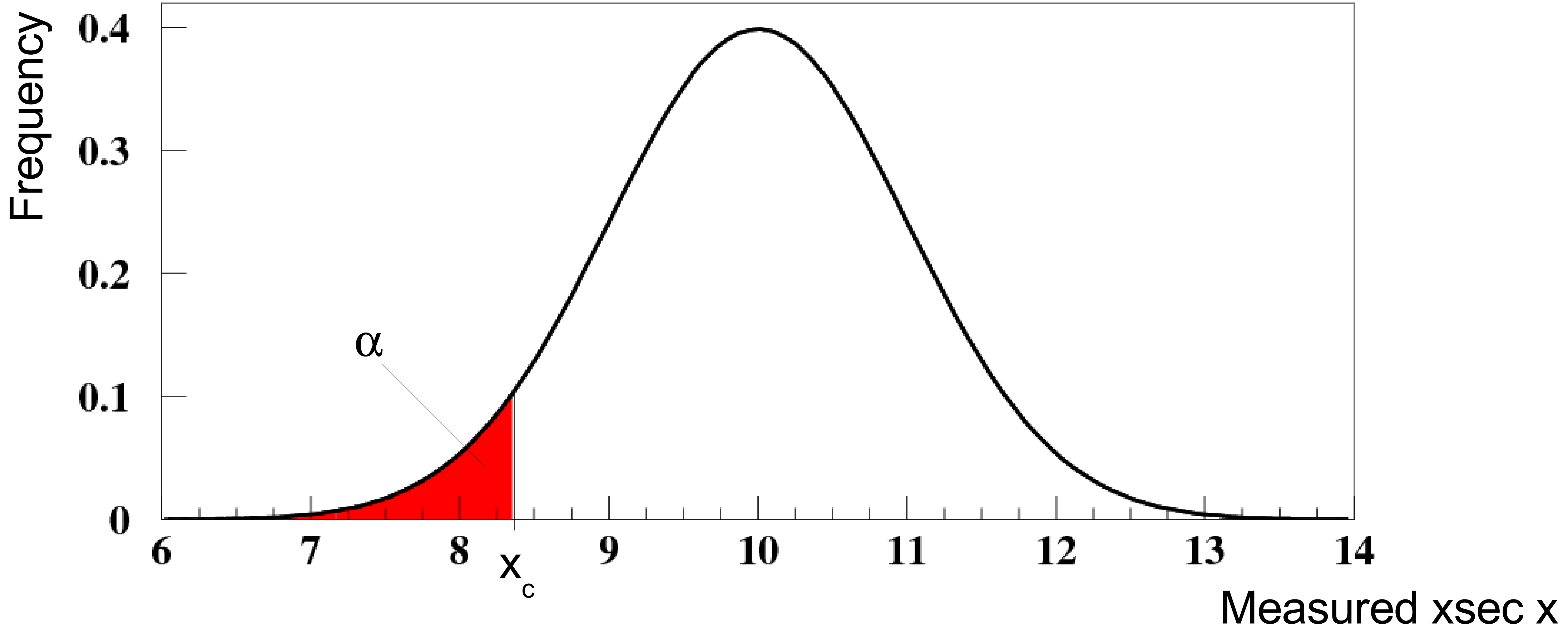
# Neyman Band for Gaussian (95% CL Upper limit)

- plot $[x_l, \infty]_\mu$ vs $\mu$ with 95% prob. x is in range; for x: max $\mu \rightarrow \mu_{limit}$



# Upper Limit and Hypothesis testing

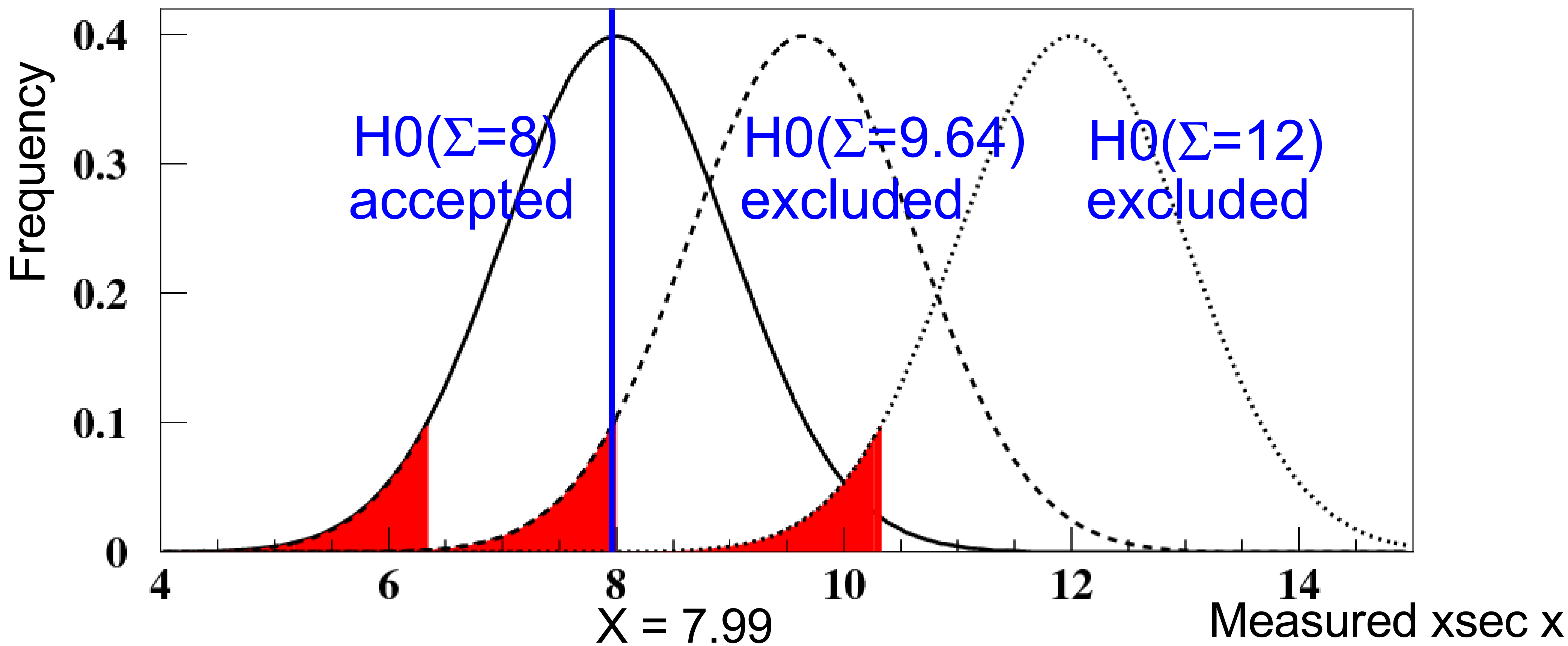
- Hypo: BSM enhanced tt production cross section $\Sigma$ is 10 (a.u.)
- Simplifying assumption: it is measured with gauss. uncertainty 1




- If $x<x_c$ this BSM model is rejected

# Upper Limit and Hypothesis testing

- $\Sigma$ is now a free parameter, determined from measurement x

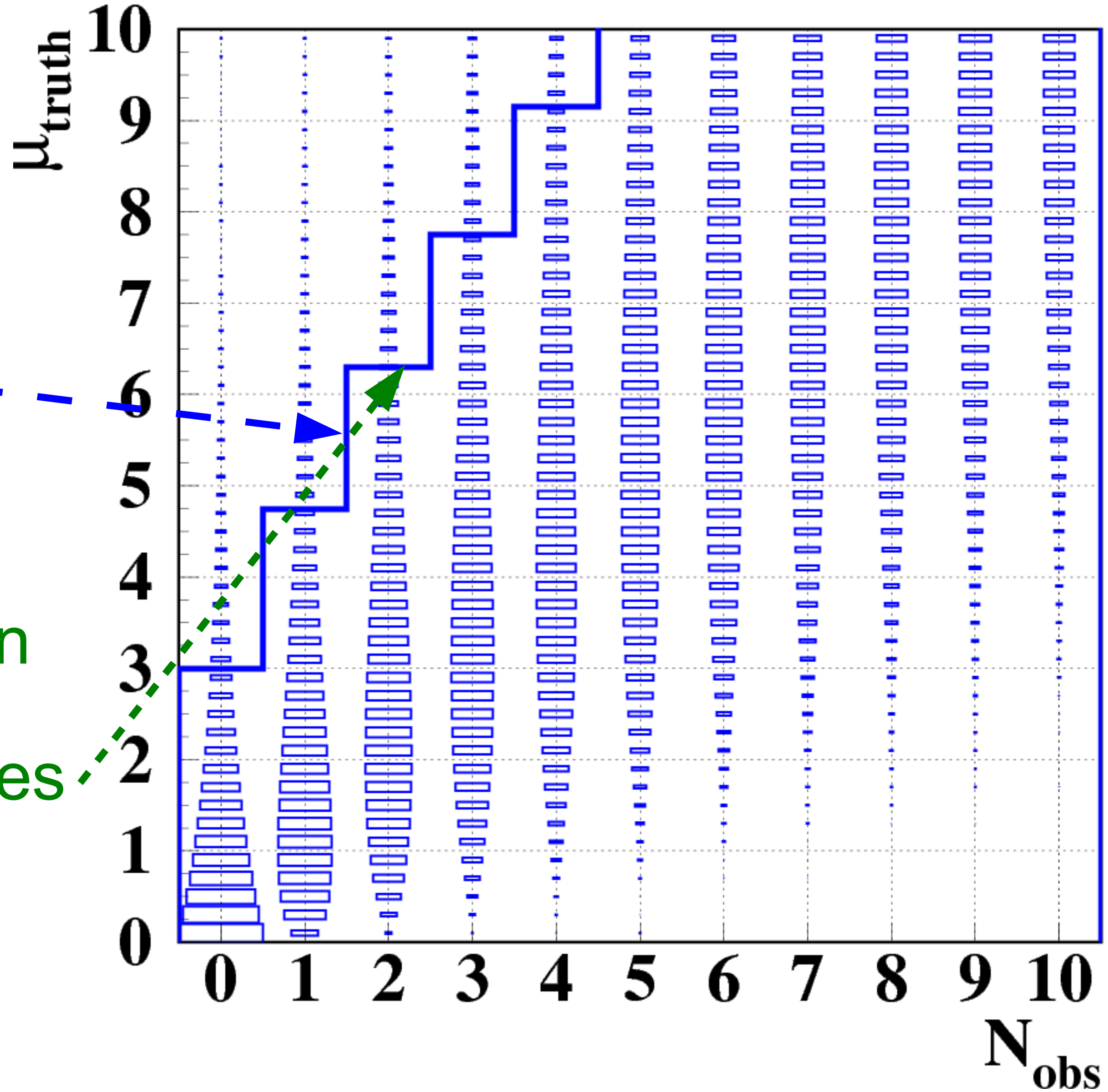



- Exclude all hypotheses with: p-value $\leq \alpha = 5\% \Leftrightarrow \Sigma > \Sigma_{\text{limit }|@95\%\text{ CL}}$
- Upper limit and signal hypothesis exclusion intrinsically linked

# 2) Poisson without & with background, expected limit, $CL_s$

# Frequentist upper limit, Poisson data

- Neyman construction: scan $\mu$ and find for each $\mu$ the max. $N_{obs}$ with $P(N \in [N_{obs}, \infty]) \geq 95\%$ $\rightarrow$ vertical lines




- For given $N_{obs}$ find largest $\mu$, where $N_{obs}$ is just contained in the interval $\rightarrow$ horizontal lines $\rightarrow$ $\mu_{limit}$

- Note: discrete $N_{obs}$ but continuous m $\rightarrow$ limit steps

# Exercise 1 (Neyman construction)

- Poisson experiment, determine limits on the parameter μ, given $N_{obs}$

  a) determine the probability α for observing a value $N<N_{obs}$ for some selected values of μ

  b) determine the limit on μ for $N_{obs}$=0,2,10,100

- Hints: the probability to find $N<N_{obs}$ is given by:

Probability: $\sum_{N=0}^{N_{\text{obs}}-1} \frac{e^{-\mu}(\mu)^N}{N!} = \alpha = \text{TMath::Prob}(2*\mu, 2*N_{\text{obs}})$

Inverse function: $\mu_{limit} = \text{TMath::ChisquareQuantile}(1-\alpha, 2*(N_{\text{obs}}+1))/2$

(a)

| μ | $N_{obs}$ | α |
|---|---|---|
| 2 | 1 | |
| 3 | 1 | |
| 5 | 2 | |
| 10 | 5 | |

(b)

| $N_{obs}$ | $\mu_{limit}$ |
|---|---|
| 0 | |
| 2 | |
| 10 | |
| 100 | |

# Exercise 1 (Neyman construction)

- Poisson experiment, determine limits on the parameter μ, given $N_{obs}$

  a) determine the probability α for observing a value $N<N_{obs}$ for some selected values of μ

  b) determine the limit on μ for $N_{obs}$=0,2,10,100

- Hints: the probability to find $N<N_{obs}$ is given by:

$$\text{Probability: } \sum_{N=0}^{N_{\text{obs}}-1} \frac{e^{-\mu}(\mu)^N}{N!} = \alpha = \text{TMath::Prob}(2*\mu, 2*N_{\text{obs}})$$

$$\text{Inverse function: } \mu_{limit} = \text{TMath::ChisquareQuantile}(1-\alpha, 2*(N_{\text{obs}}+1))/2$$

(a)

| μ | $N_{obs}$ | α |
|---|---|---|
| 2 | 1 | 0.14 |
| 3 | 1 | 0.05 |
| 5 | 2 | 0.04 |
| 10 | 5 | 0.03 |

(b)

| $N_{obs}$ | $\mu_{limit}$ |
|---|---|
| 0 | 3.0 |
| 2 | 6.3 |
| 10 | 17.0 |
| 100 | 118.1 |

# Exercise 1 (Neyman construction)

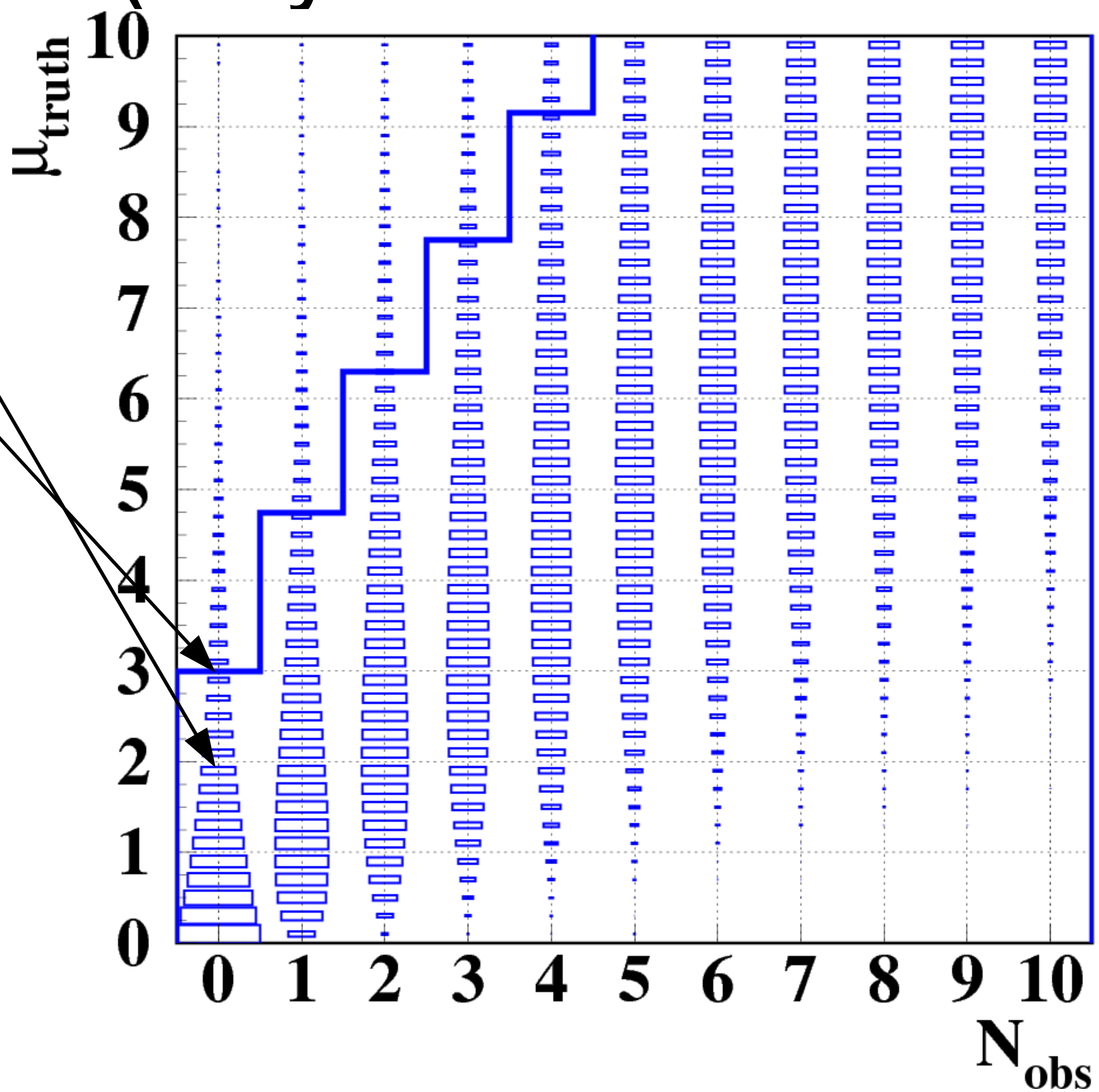
(a)

P(N<N$_{obs}$)

| μ | $N_{obs}$ | α |
|---|---|---|
| 2 | 1 | 0.14 |
| 3 | 1 | 0.05 |
| 5 | 2 | 0.04 |
| 10 | 5 | 0.03 |

(b)

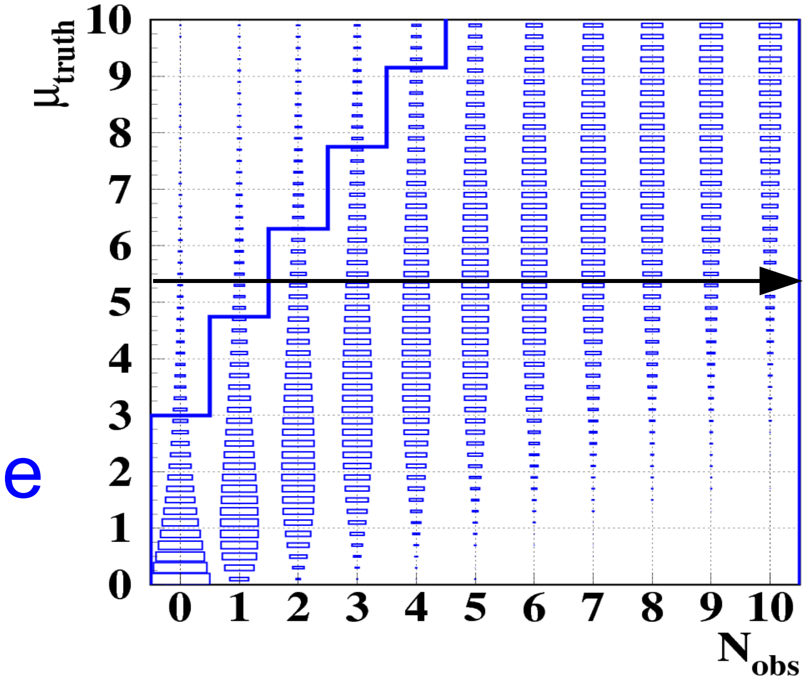
| $N_{obs}$ | $\mu_{limit}$ |
|---|---|
| 0 | 3.0 |
| 2 | 6.3 |
| 10 | 17.0 |
| 100 | 118.1 |




Probability: $\sum_{N=0}^{N_{obs}-1} \frac{e^{-\mu}(\mu)^N}{N!} = \alpha = \text{TMath::Prob}(2*\mu, 2*N_{obs})$

Inverse function: $\mu_{limit} = \text{TMath::ChisquareQuantile}(1-\alpha, 2*(N_{obs}+1))/2$

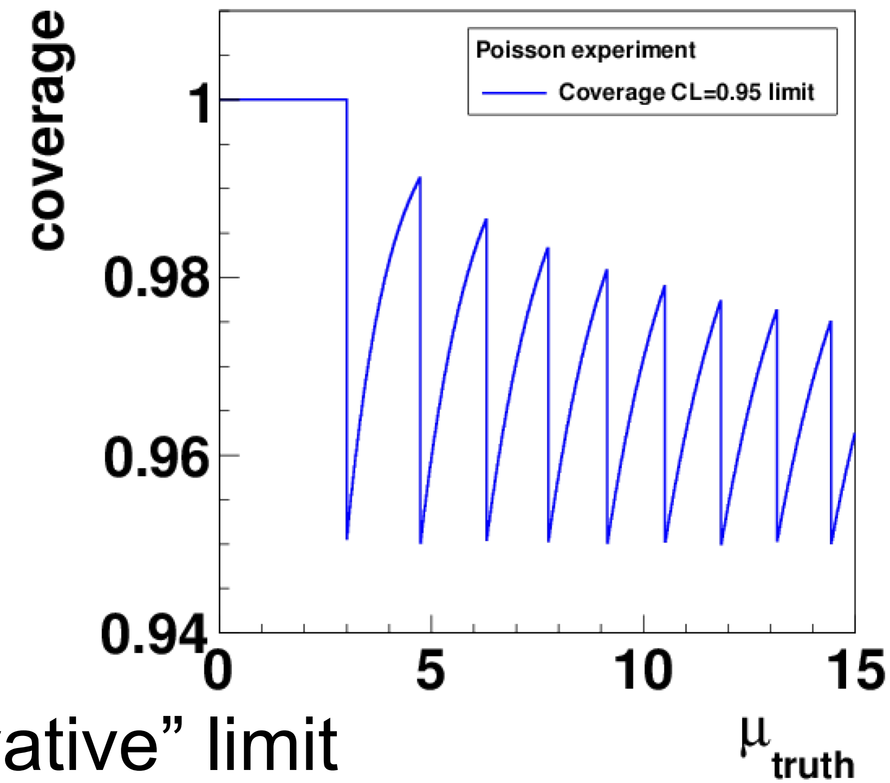
# Coverage




- Coverage: given the limit procedure, calculate for each $\mu_{truth}$ probability to include the true value in the Confidence interval
- Poisson example (exercise 2)

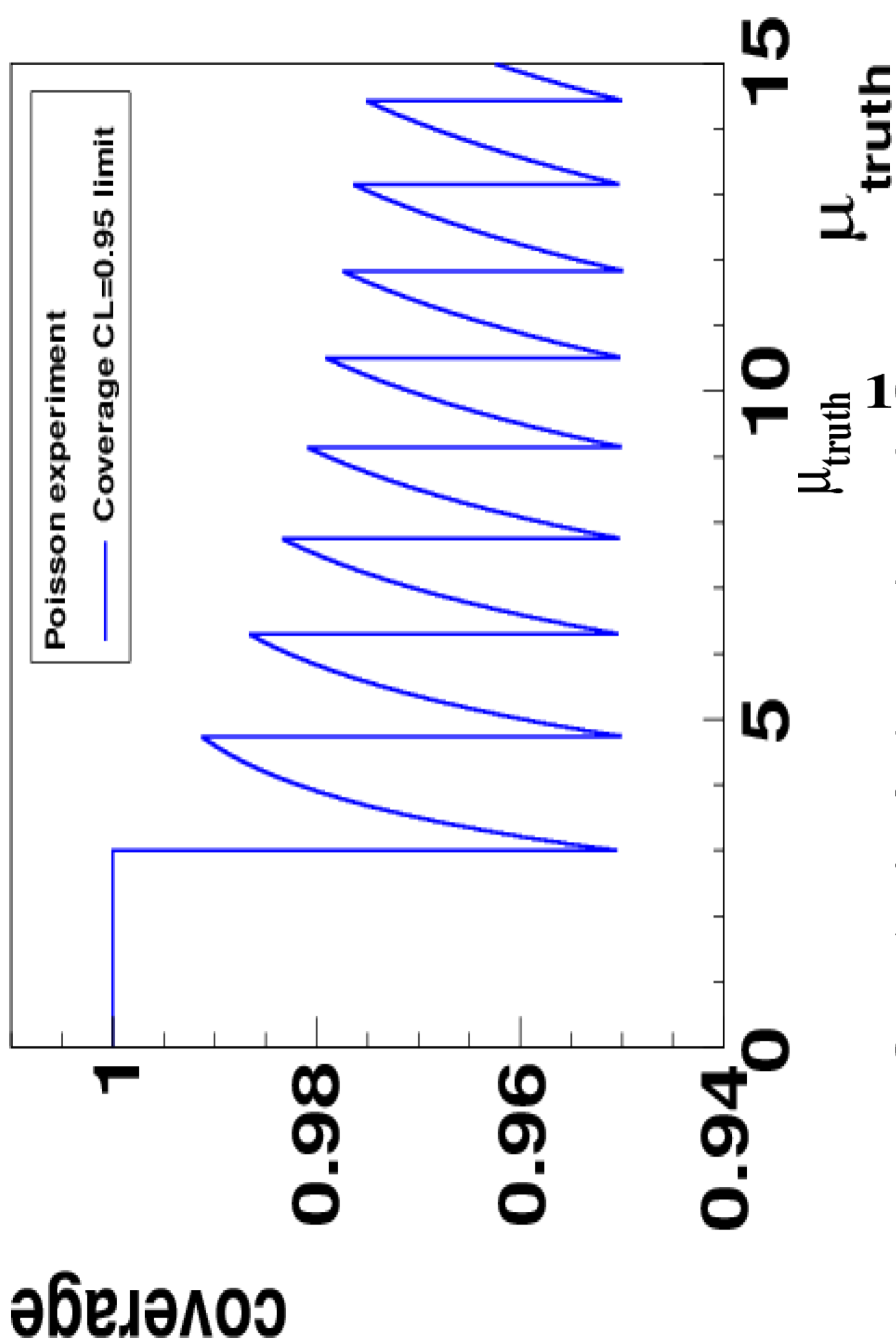
$$P_{incl}(\mu_{truth}) = \sum_N P_{\mu,\, truth}(N)\,\Theta(\mu_{truth} \leq \mu_{limit}(N))$$

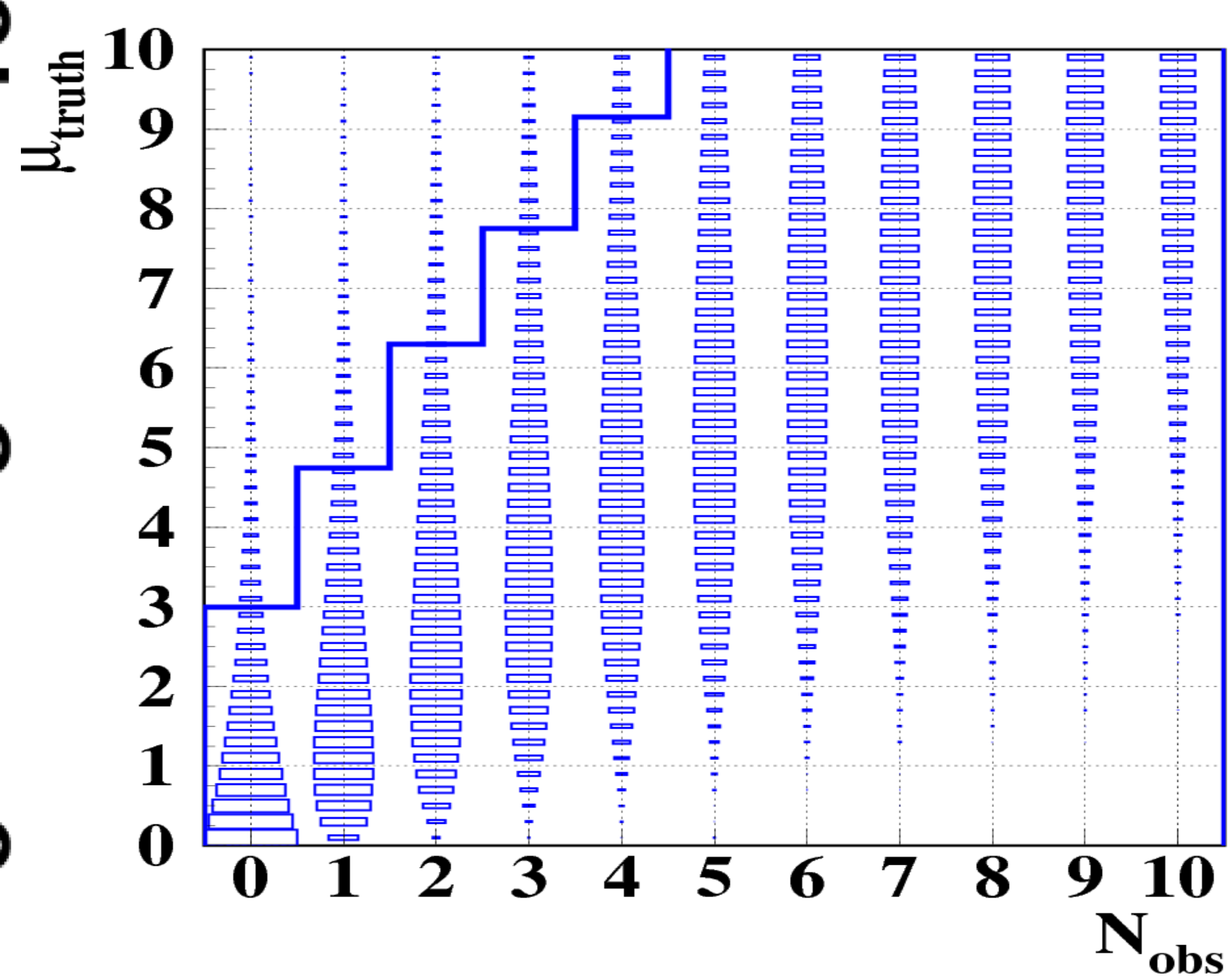
$$\text{where } \Theta(\mu_{truth} \leq \mu_{limit}) = \begin{cases} 1 & \text{if } \mu_{truth} \leq \mu_{limit} \\ 0 & \text{otherwise} \end{cases}$$




- coverage=0.95: exact coverage
- coverage<0.95: undercoverage
- coverage>0.95: overcoverage, “conservative” limit
- “Simple” Poisson case: overcoverage (discrete measurement)

# Coverage illustration



# Bayesian upper limits

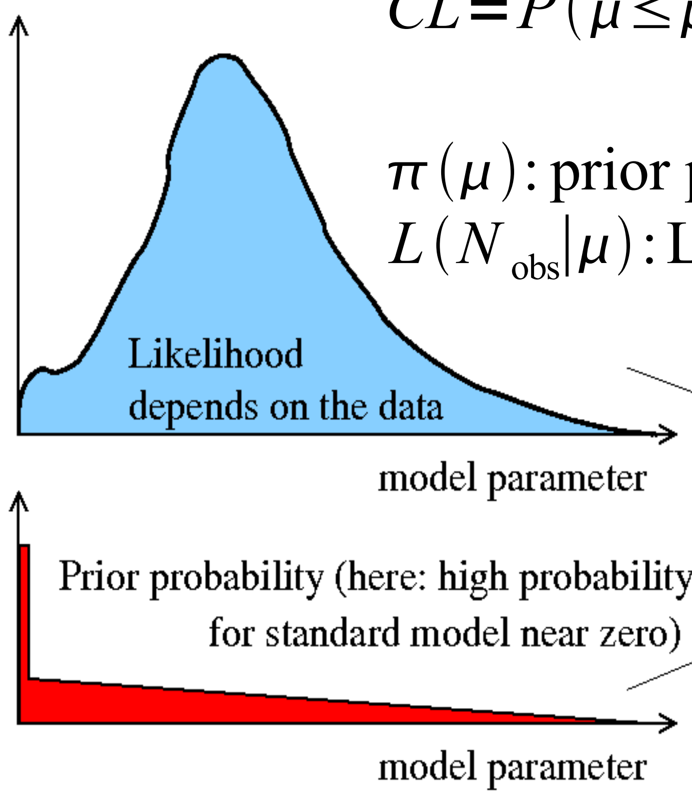
- Bayesian limit: exclude a set of theories, such that the posterior probability of the excluded theories is 1-CL

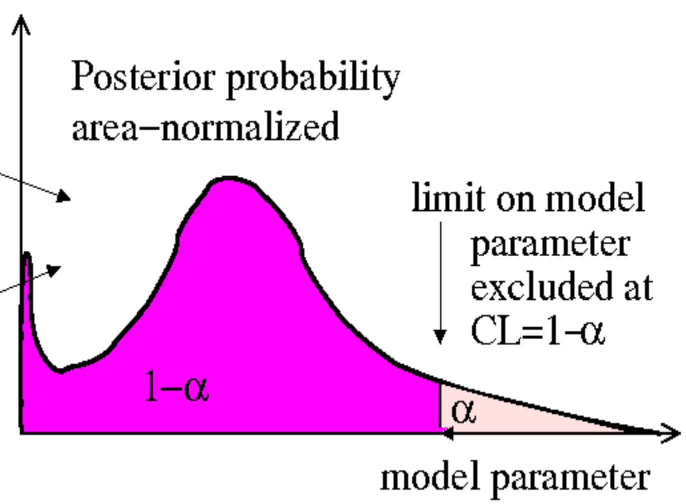
Enumerator: integrate over allowed theories

$$CL = P(\mu \leq \mu_{limit} | N_{obs}) = 1 - \alpha = \frac{\int_0^{\mu_{limit}} L(N_{obs}|\mu)\pi(\mu)\,d\mu}{\int_0^{\infty} L(N_{obs}|\mu)\pi(\mu)\,d\mu}$$

$\pi(\mu)$: prior probability of the model $\mu$
$L(N_{obs}|\mu)$: Likelihood

Denominator: integrate all theories (normalisation)




Bayesian limit: integrate over models, fixed $N_{obs}$

Frequentist limit: integrate over $N_{obs}$, test each model

# Exercise 2 (Bayesian limit)

- Exercise 2a: Bayesian limit for $N_{obs}$=0,2,10,100 (flat prior)

  (use Root macro)

- Exercise 2b: use a prior P(μ)=μ, $N_{obs}$={0,2,10,100}

  (modify first routine in macro)

- Compare to exercise 2

- Bayesian limit with arbitrary prior → numerical integration

- `GetPosterior.C(muLimit,nObs)`

$$\text{Posterior} \sim \int_0^{\mu_0} d\mu \, \text{Prior}(\mu) \frac{\exp[-\mu]\mu^{N_{obs}}}{N_{obs}!}$$

- Vary muLimit until Posterior=0.95

| | frequentist | Bayes flat | Bayes P(μ)=μ |
|---|---|---|---|
| $N_{obs}$ | $\mu_{limit}$ | $\mu_{limit}$ | $\mu_{limit}$ |
| 0 | 3.0 | | |
| 2 | 6.3 | | |
| 10 | 17.0 | | |
| 100 | 118.1 | | |

# Bayesian limit exercise

- Exercise 2a: Bayesian limit for $N_{obs}$=0,2,10,100 (flat prior) (use Root macro)
- Exercise 2b: use a prior P(μ)=μ, $N_{obs}$={0,2,10,100}

- For this example: Bayes flat=Frequentist
- Prior P(μ)=μ gives more conservative limit

| | frequentist | Bayes flat | Bayes P(μ)=μ |
|---|---|---|---|
| $N_{obs}$ | $\mu_{limit}$ | $\mu_{limit}$ | $\mu_{limit}$ |
| 0 | 3.0 | 3.0 | 4.7 |
| 2 | 6.3 | 6.3 | 7.8 |
| 10 | 17.0 | 17.0 | 18.2 |
| 100 | 118.1 | 118.2 | 119.3 |

# Limits with background

- Expected number of events:

  $\mu = s + b$, $s$, $b$: signal and background event yield, respectively

  - s=0: standard model
  - s>0: new physics

- Assume background known. What is the limit on the signal?
- Frequentist: set limit on μ, then subtract b
- Bayesian: use prior probability which is zero for s<0

# Exercise 3 (limit with background)

- Calculate Frequentist and Bayesian limits for $N_{obs}$={0,2} and b={0.5,2.0,3.5}

Poisson parameter: $\mu = s + b$

| | b=0.5 | | b=2.0 | | b=3.5 | |
|---|---|---|---|---|---|---|
| | $N_{obs}$=0 | $N_{obs}$=2 | $N_{obs}$=0 | $N_{obs}$=2 | $N_{obs}$=0 | $N_{obs}$=2 |
| Bayesian | | | | | | |
| Frequentist | | | | | | |

- Frequentist: use methods from exercise 2
- use macro GetPosteriorWithBackground.C

# Exercise 3 (limit with background)

- Calculate Frequentist and Bayesian limits for $N_{obs}$={0,2} and b={0.5,2.0,3.5}

Poisson parameter: $\mu = s + b$

| | b=0.5 | | b=2.0 | | b=3.5 | |
|---|---|---|---|---|---|---|
| | $N_{obs}$=0 | $N_{obs}$=2 | $N_{obs}$=0 | $N_{obs}$=2 | $N_{obs}$=0 | $N_{obs}$=2 |
| Bayesian | 3.0 | 5.8 | 3.0 | 4.8 | 3.0 | 4.3 |
| Frequentist | 2.5 | 5.8 | 1.0 | 4.3 | -0.5 | 2.8 |

- Problem for Frequentist limit, $N_{obs}$=0 and b=3.5:

limit excludes all signal above s=-0.5.

Even the “standard model” s=0 is excluded

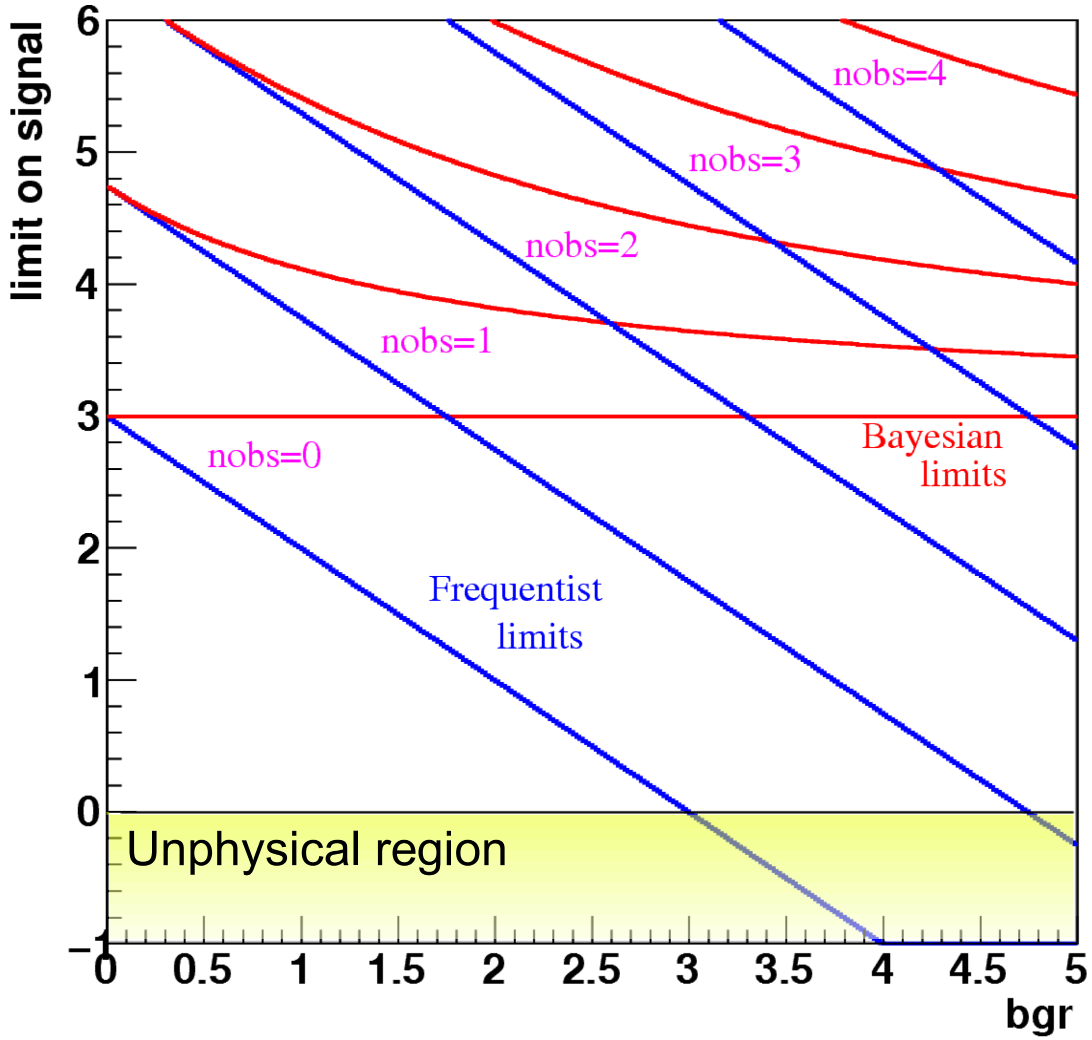
# Discussion Exercise 3

- Frequentist analysis can give limits where all models are “excluded” at a given CL (even model with s=0)

  $N_{obs}$=0, μ=s+b, b=3.5

  → limit s<-0.5 @ 95% CL but s>=0 physical bound

- Bayesian limit uses prior knowledge s>=0




- Feldman-Cousins frequentist approach based on likelihood ratio test statistics provides an alternative (see later)

# Limits near a boundary

- What to do if frequentist analysis excludes parameters beyond the sensitivity of the experiment or beyond boundaries?
- Give also expected limit to show sensitivity of the experiment (exercise 4)
- $CL_s$ method, also known as “modified frequentist” (exercise 5)
- Bayesian methods (see exercise 3)

# Expected limit (exercise 4)

- Expected limit: limit weighted by background probability

$$\langle s_{\text{limit}} \rangle = \sum_{n=0}^{\infty} \frac{e^{-b} b^n}{n!} \text{LimitOnSignal}(b, n)$$

| | b=0.5 | | b=2.0 | | b=3.5 | |
|---|---|---|---|---|---|---|
| | $N_{obs}$=0 | $N_{obs}$=2 | $N_{obs}$=0 | $N_{obs}$=2 | $N_{obs}$=0 | $N_{obs}$=2 |
| Bayesian | 3.0 | 5.8 | 3.0 | 4.8 | 3.0 | 4.3 |
| Frequentist | 2.5 | 5.8 | 1.0 | 4.3 | -0.5 | 2.8 |
| Expected | | | | | | |

- Calculate expected limits for b={0.5,2.0,3.5}
- Macro GetExpectedLimit.C

# Expected limit (exercise 4)

- Expected limit: limit weighted by background probability

$$\langle s_{\text{limit}} \rangle = \sum_{n=0}^{\infty} \frac{e^{-b} b^n}{n!} \text{LimitOnSignal}(b, n)$$

| | b=0.5 | | b=2.0 | | b=3.5 | |
|---|---|---|---|---|---|---|
| | $N_{obs}$=0 | $N_{obs}$=2 | $N_{obs}$=0 | $N_{obs}$=2 | $N_{obs}$=0 | $N_{obs}$=2 |
| Bayesian | 3.0 | 5.8 | 3.0 | 4.8 | 3.0 | 4.3 |
| Frequentist | 2.5 | 5.8 | 1.0 | 4.3 | -0.5 | 2.8 |
| Expected | 3.3 | | 4.2 | | 4.9 | |

- Problematic case: expected limit differs a lot from observed limit

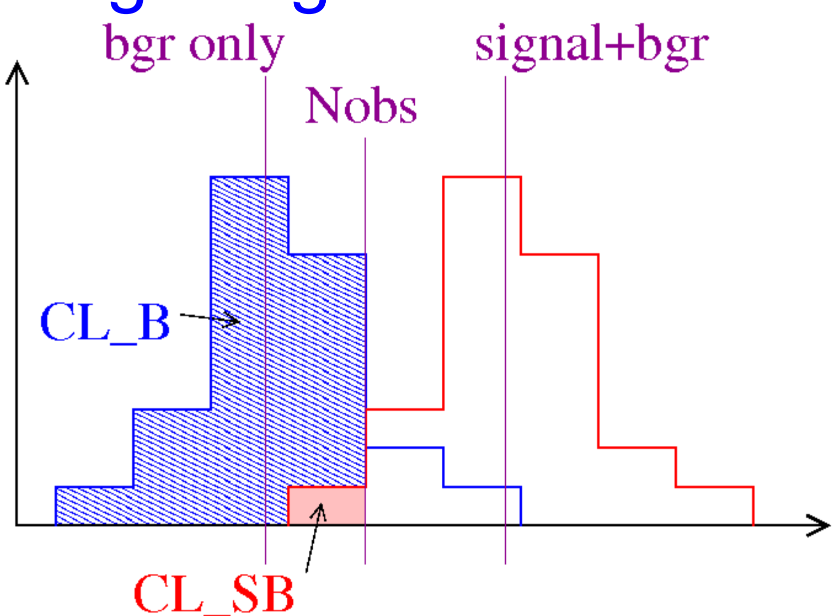
→ Recognize statistical fluctuation or problem with background

# The $CL_S$ (modified frequentist) method

- Frequentist limit: $$1 - CL \geq \alpha = \mathrm{CL}_{SB} = P\left(N \leq N_{\mathrm{obs}}; \mu = s + b\right)$$
- $CL_S$ limit: $$1 - CL \geq \mathrm{CL}_S = \frac{\mathrm{CL}_{SB}}{\mathrm{CL}_B} = \frac{P\left(N \leq N_{\mathrm{obs}}; \mu = s + b\right)}{P\left(N \leq N_{\mathrm{obs}}; \mu = b\right)}$$
- Probability is normalized to background probability
- $CL_B \leq 1 \rightarrow CL_S \geq CL_{SB}$: same α requires larger signal

  Limit is "conservative"
- For zero signal: $CL_S=1$

  → zero signal is never excluded



# Exercise 5 ($CL_s$ method)

- Frequentist limit: $1 - CL \geq \alpha = \mathrm{CL}_{SB} = P(N \leq N_{\mathrm{obs}}; \mu = s + b)$
- $CL_s$ limit: $1 - CL \geq \mathrm{CL}_S = \frac{\mathrm{CL}_{SB}}{\mathrm{CL}_B} = \frac{P(N \leq N_{\mathrm{obs}}; \mu = s + b)}{P(N \leq N_{\mathrm{obs}}; \mu = b)}$

| | b=0.5 | | b=2.0 | | b=3.5 | |
|---|---|---|---|---|---|---|
| | $N_{obs}$=0 | $N_{obs}$=2 | $N_{obs}$=0 | $N_{obs}$=2 | $N_{obs}$=0 | $N_{obs}$=2 |
| Bayesian | 3.0 | 5.8 | 3.0 | 4.8 | 3.0 | 4.3 |
| Frequentist | 2.5 | 5.8 | 1.0 | 4.3 | -0.5 | 2.8 |
| $CL_s$ | | | | | | |
| Expected | 3.3 | | 4.2 | | 4.9 | |

- Use macro GetCLsLimit.C to calculate $CL_s$, iterate to get limit

# Exercise 5 ($CL_s$ method)

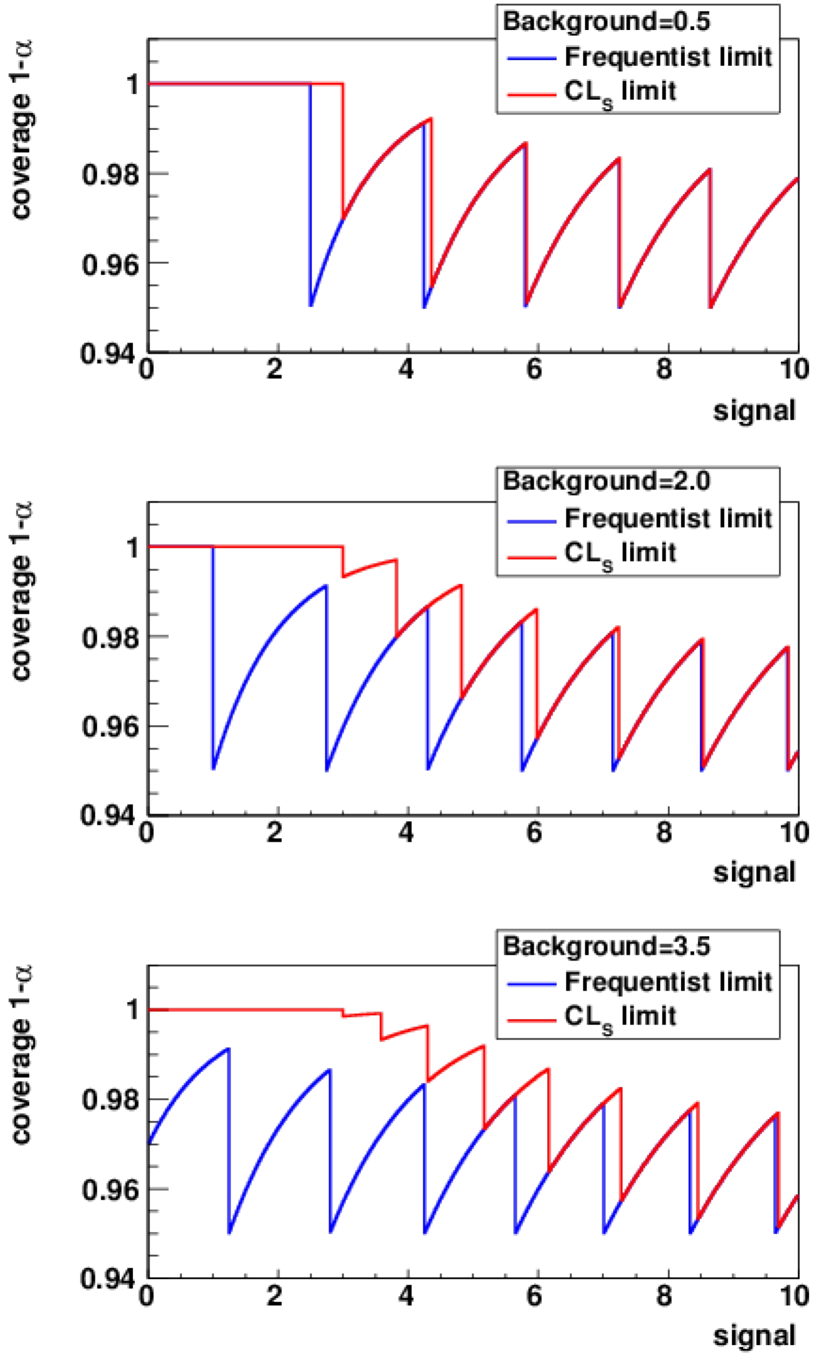
- Frequentist limit: $1-CL \geq \alpha = \mathrm{CL}_{SB} = P(N \leq N_{\mathrm{obs}}; \mu = s+b)$
- $CL_s$ limit: $1-CL \geq \mathrm{CL}_S = \dfrac{\mathrm{CL}_{SB}}{\mathrm{CL}_B} = \dfrac{P(N \leq N_{\mathrm{obs}}; \mu = s+b)}{P(N \leq N_{\mathrm{obs}}; \mu = b)}$

| | b=0.5 | | b=2.0 | | b=3.5 | |
|---|---|---|---|---|---|---|
| | $N_{obs}=0$ | $N_{obs}=2$ | $N_{obs}=0$ | $N_{obs}=2$ | $N_{obs}=0$ | $N_{obs}=2$ |
| Bayesian | 3.0 | 5.8 | 3.0 | 4.8 | 3.0 | 4.3 |
| Frequentist | 2.5 | 5.8 | 1.0 | 4.3 | -0.5 | 2.8 |
| $CL_s$ | 3.0 | 5.8 | 3.0 | 4.8 | 3.0 | 4.3 |
| Expected | 3.3 | | 4.2 | | 4.9 | |

- For this example, $CL_s$ is identical to Bayesian (with flat prior)

# Summary of CLs pros & cons

- $CL_s$ method avoids problem with limits better than the experiments sensitivity
- Limits on s always > 0
- Disadvantage: $CL_s$ method is conservative, in particular for small signals

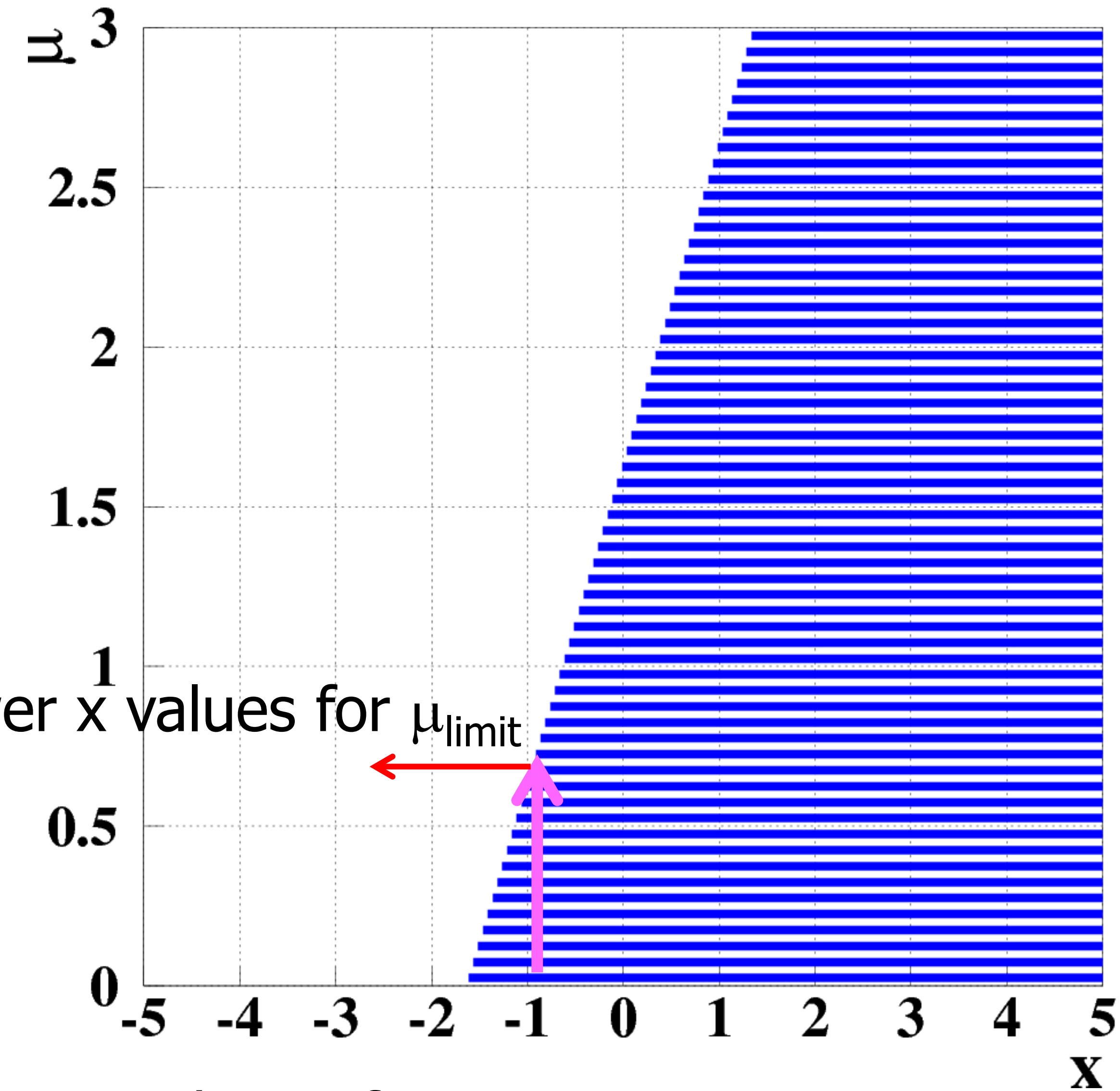


# 3) Likelihood ratio based intervals

# Upper limit for gaussian with boundary $\mu \geq 0$

Measurement x of $\mu$ with gaussian uncert.

$f(x|\mu) \sim \exp(-(x-\mu)^2/2)$

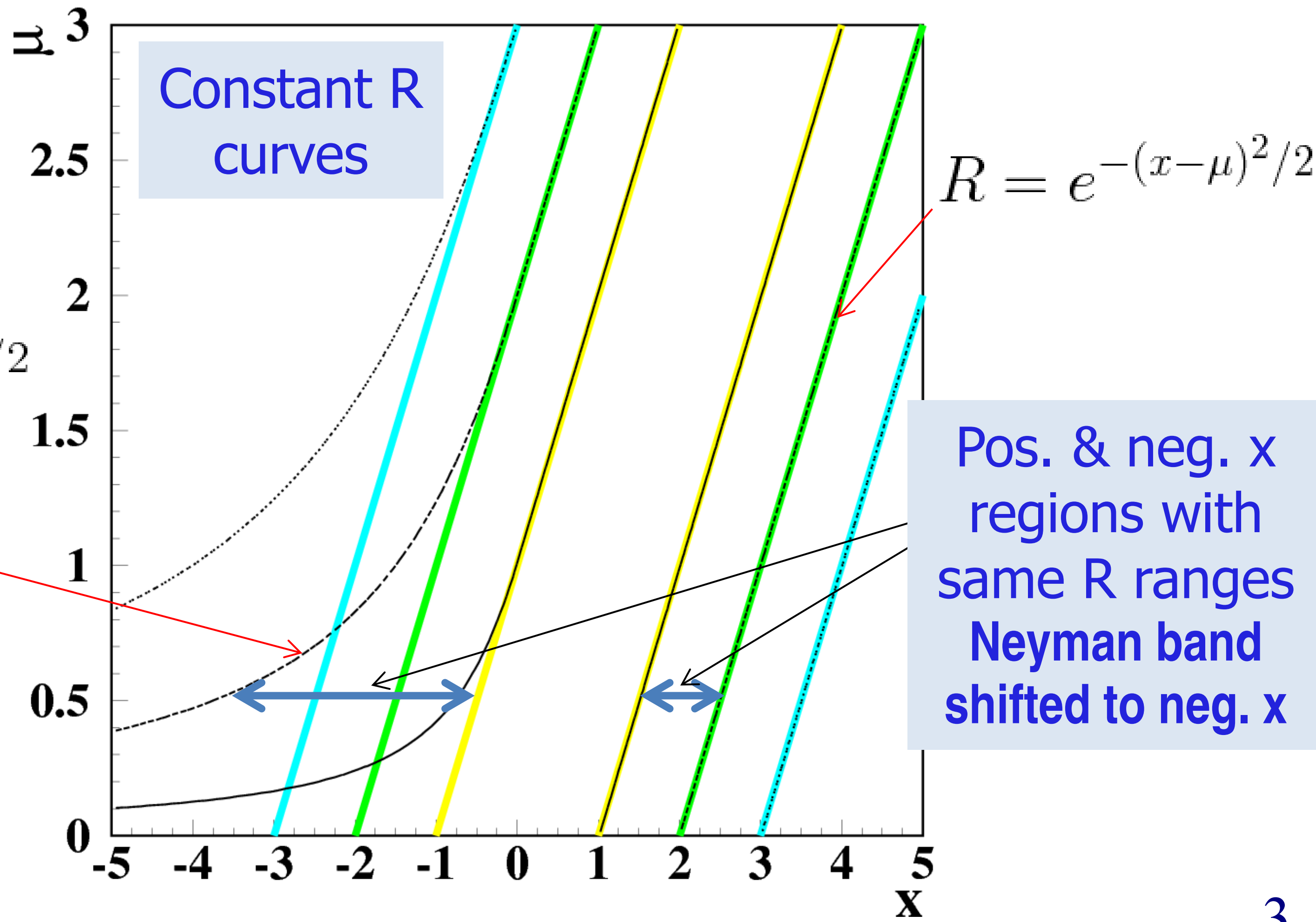



- Problem: upper limit approaches 0 for x going negative, rule out all positive $\mu$ values for $x<-1.64$ !

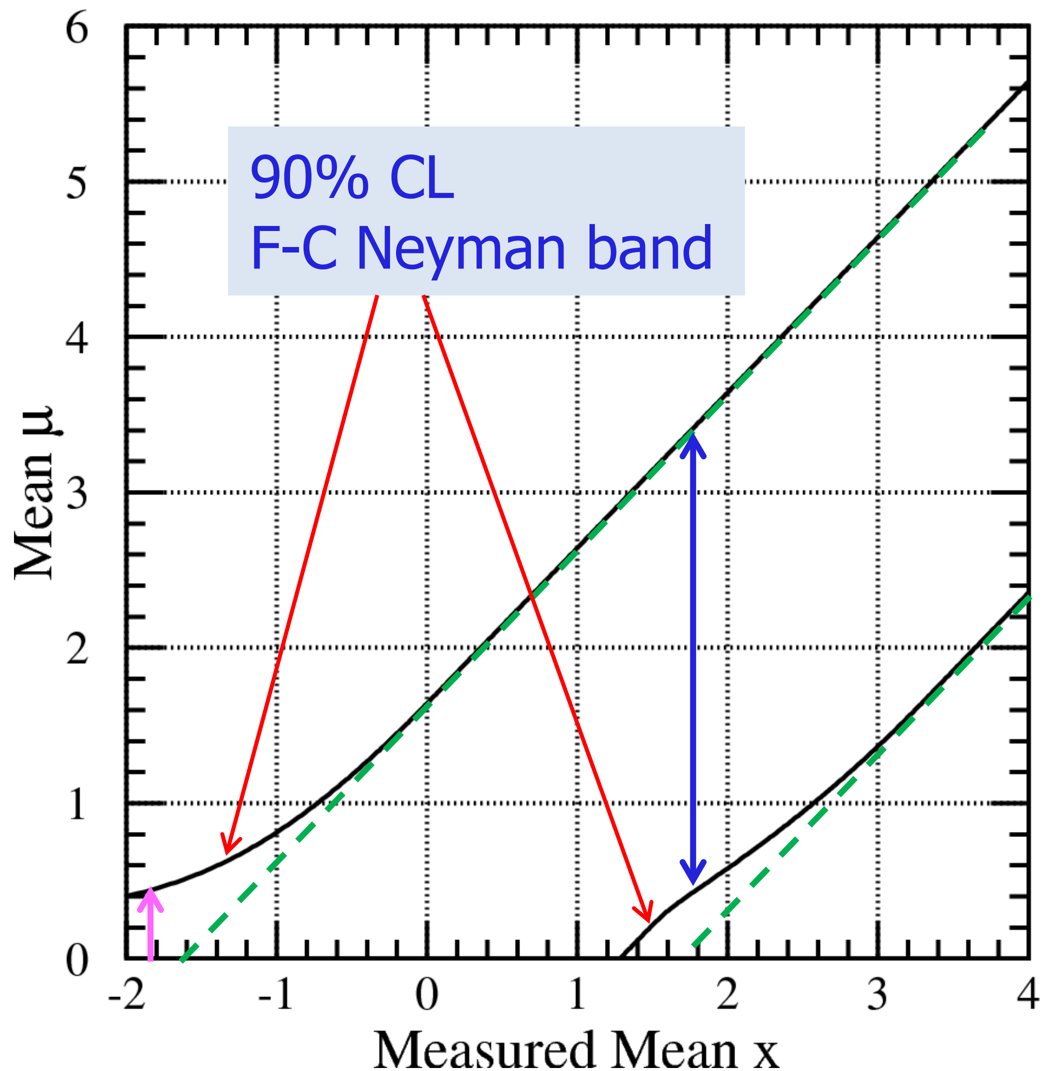
# Feldman-Cousins unified approach Phys.Rev.D57:387303889,1998

- Use different ordering principle for test statistics t: so far: t=x
- $t = R = \dfrac{L(x|\mu)}{L(x|\mu_{best})}$ with $\mu_{best}$ = best fit value in physical range:
- $x \geq 0$: $\mu_{best} = x$
- $x<0$: $\mu_{best} = 0$
- $\mu_{best} = \max(x,0)$

$$R = e^{-(x-\mu)^2/2}/e^{-x^2/2} = e^{x\mu-\mu^2}$$

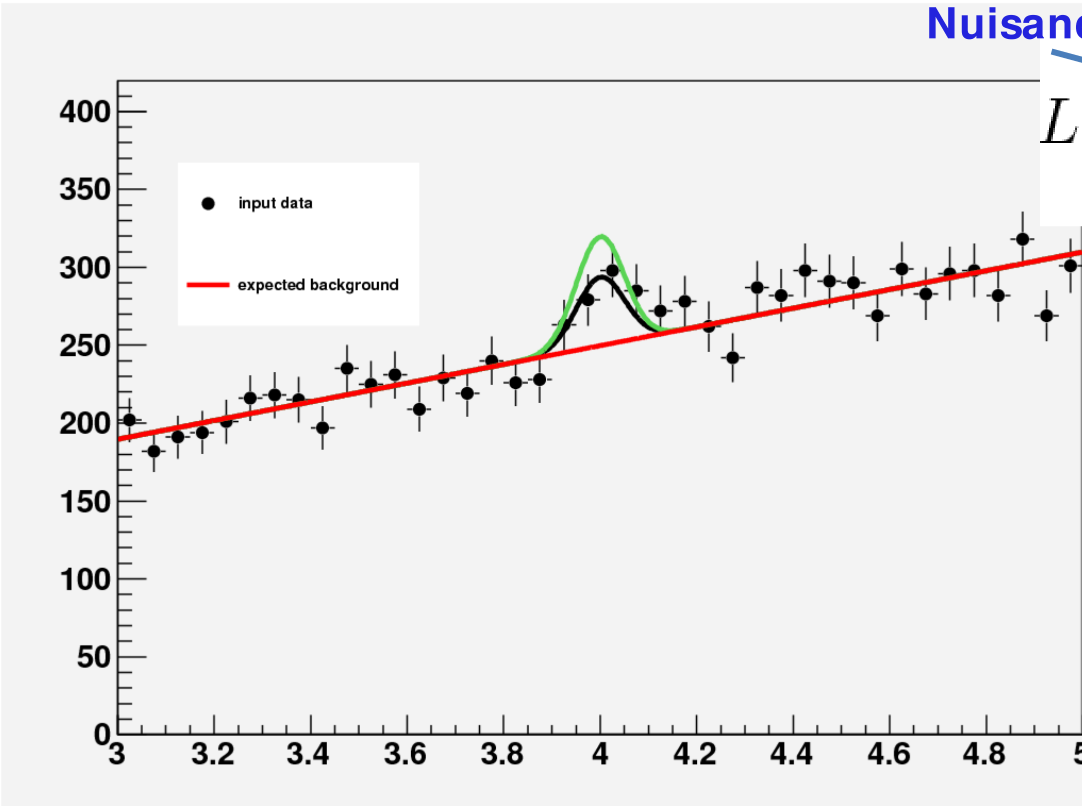


# Feldman-Cousins unified approach Phys.Rev.D57:387303889,1998




- Note: F-C switches itself as function of x from upper limit to two-sided interval → "unified approach"

# Searches using Likelihood ratio




Nuisance pars

$$L(\mu, \boldsymbol{\theta}) = \prod_{j=1}^{N} \frac{(\mu s_j + b_j)^{n_j}}{n_j!} e^{-(\mu s_j + b_j)}$$

·L(control data)·PDFs(θ)

For testing $\mu$:

$$\lambda(\mu) = \frac{L(\mu, \hat{\hat{\boldsymbol{\theta}}})}{L(\hat{\mu}, \hat{\boldsymbol{\theta}})}$$

Profiled values

Best fit values

Principal test statistics:

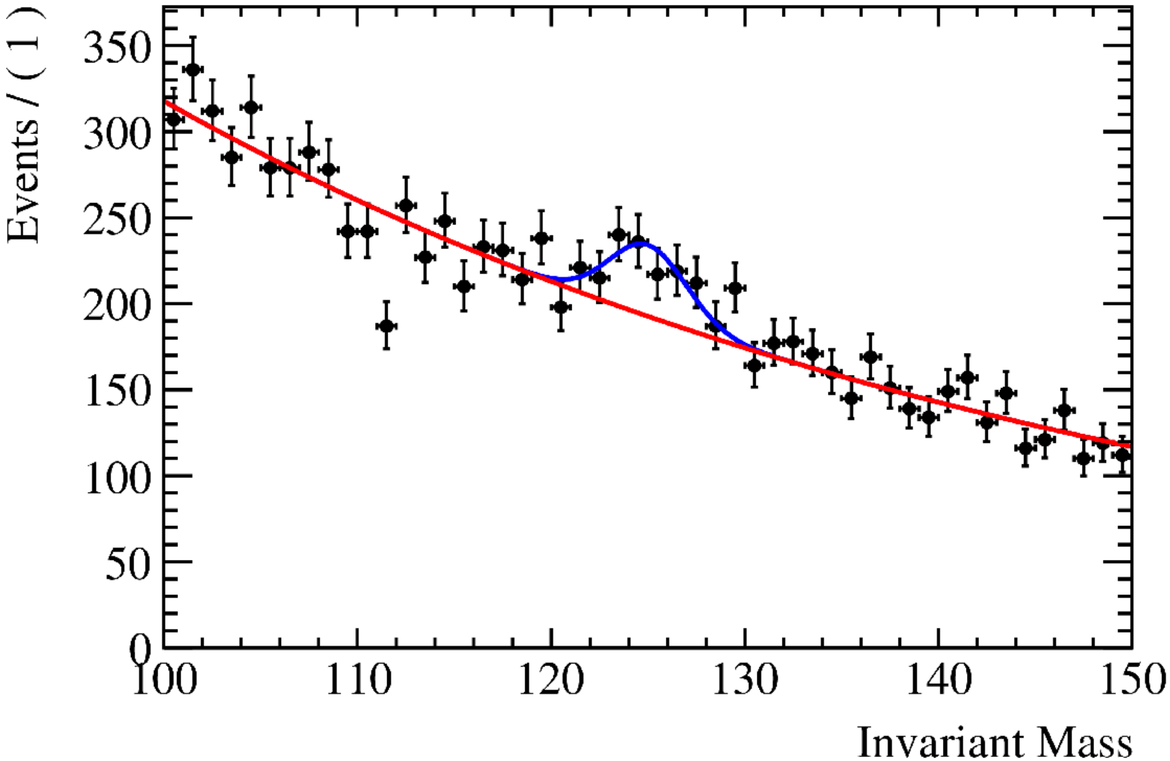
$$q_\mu = -2 \ln \lambda(\mu)$$

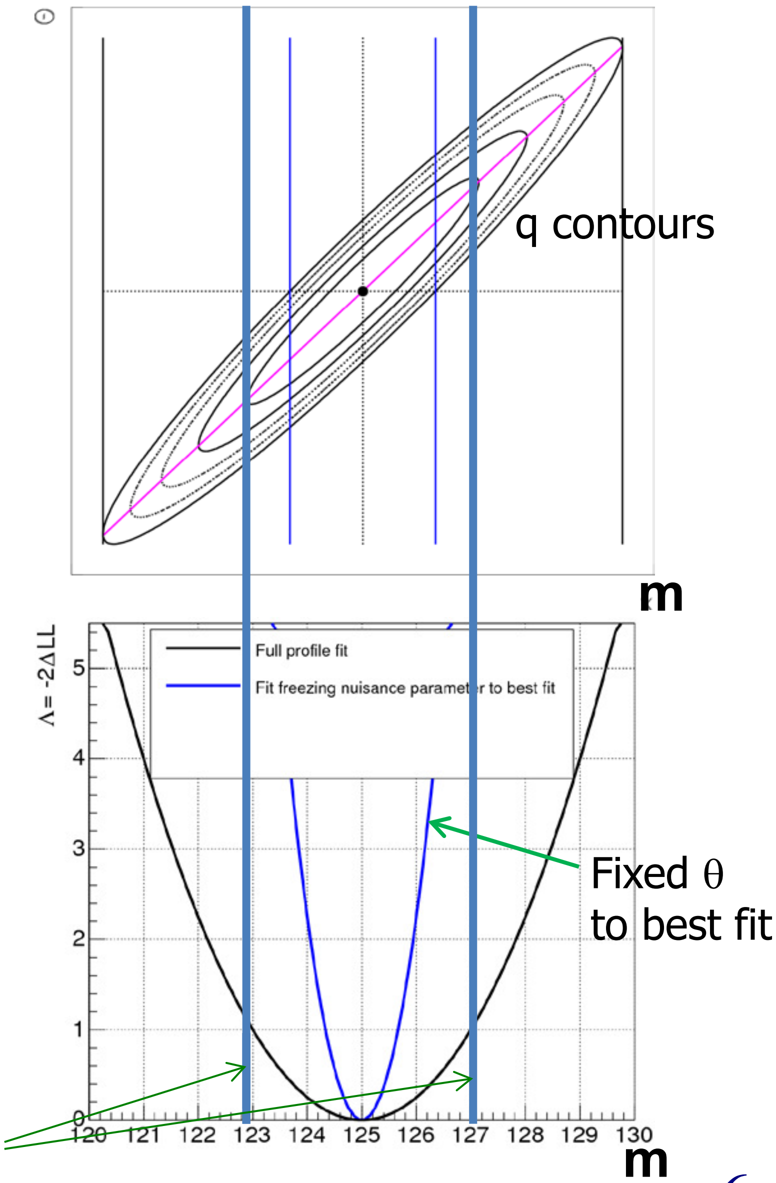
## Discuss cases:

- Clear peak+large stat.
- Small or negative bump:
  - Asymptotics (large stat.)
  - Toys (any stat.)

# Clear peak + large stat.: fit mass

Profile likelihood: Scan
$q = -2\Delta\ln(L)$ vs m whilst profiling $\theta$



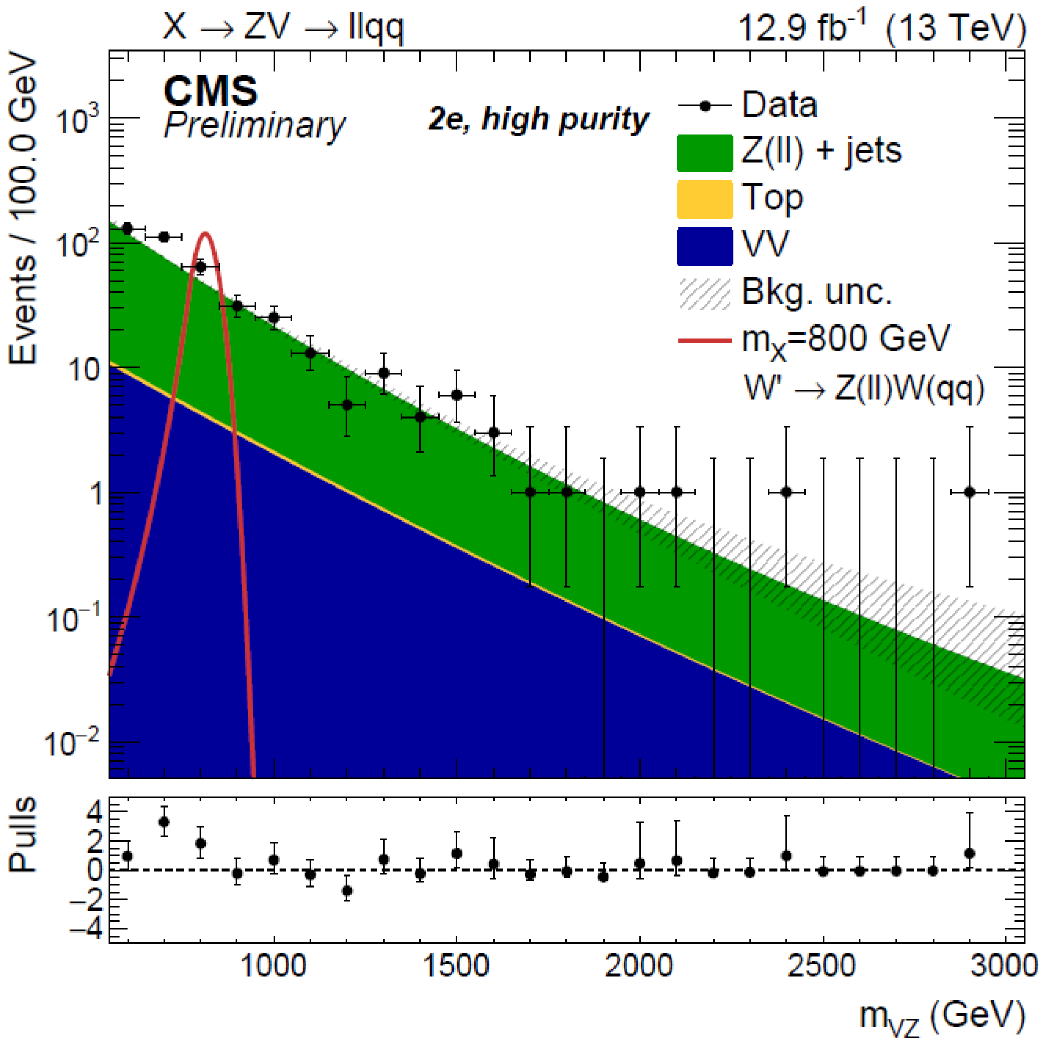

Gaussian shaped s + exponential b

Fix: $\mu=1$=signal strength; Fit: mass m and
$\theta$ = b exponential slope (nuisance)

→ 68% CL interval from m points with q=1

# Searches using Likelihood ratio

CMS B2G-16-022




Nuisance pars

$$L(\mu, \boldsymbol{\theta}) = \prod_{j=1}^{N} \frac{(\mu s_j + b_j)^{n_j}}{n_j!} e^{-(\mu s_j + b_j)}$$

·L(control data)·PDFs(θ)

$$\lambda(\mu) = \frac{L(\mu, \hat{\hat{\boldsymbol{\theta}}})}{L(\hat{\mu}, \hat{\boldsymbol{\theta}})}$$

Profiled values

Best fit values

for $\hat{\mu}$< 0 replace by

$$L(0, \hat{\hat{\boldsymbol{\theta}}}(0))$$

LHC test statistics

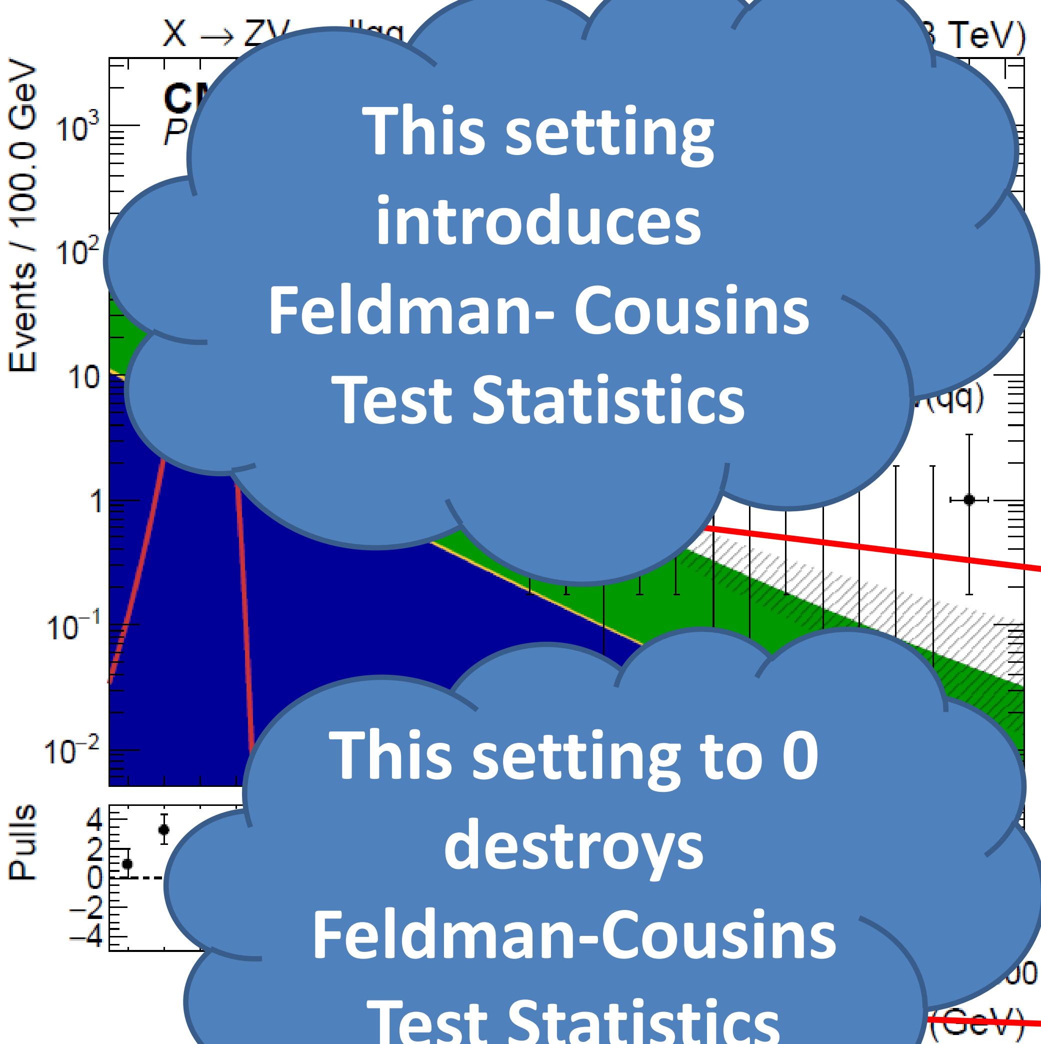
CCGV paper arXiv:1007.1727

$$q_\mu = -2 \ln \lambda(\mu)$$

| | |
|---|---|
| Discovery | $q_0$ (set to 0 for $\hat{\mu} \le 0$) |
| Upper limit | $q_\mu$ (set to 0 for $\hat{\mu} \ge \mu$) |

# Searches using Likelihood ratio

CMS B2G-16-022




This setting introduces Feldman- Cousins Test Statistics

This setting to 0 destroys Feldman-Cousins Test Statistics

Nuisance pars

$$L(\mu,\boldsymbol{\theta}) = \prod_{j=1}^{N} \frac{(\mu s_j + b_j)^{n_j}}{n_j!} e^{-(\mu s_j + b_j)}$$

$\cdot$L(control data)$\cdot$PDFs($\theta$)

Profiled values

$$\lambda(\mu) = \frac{L(\mu, \hat{\hat{\boldsymbol{\theta}}})}{L(\hat{\mu}, \hat{\boldsymbol{\theta}})}$$

Best fit values

for $\hat{\mu} < 0$ replace by $L(0, \hat{\hat{\boldsymbol{\theta}}}(0))$

LHC test statistics

CCGV paper arXiv:1007.1727

$$q_\mu = -2 \ln \lambda(\mu)$$

Discovery $q_0$ (set to 0 for $\hat{\mu} \leq 0$)

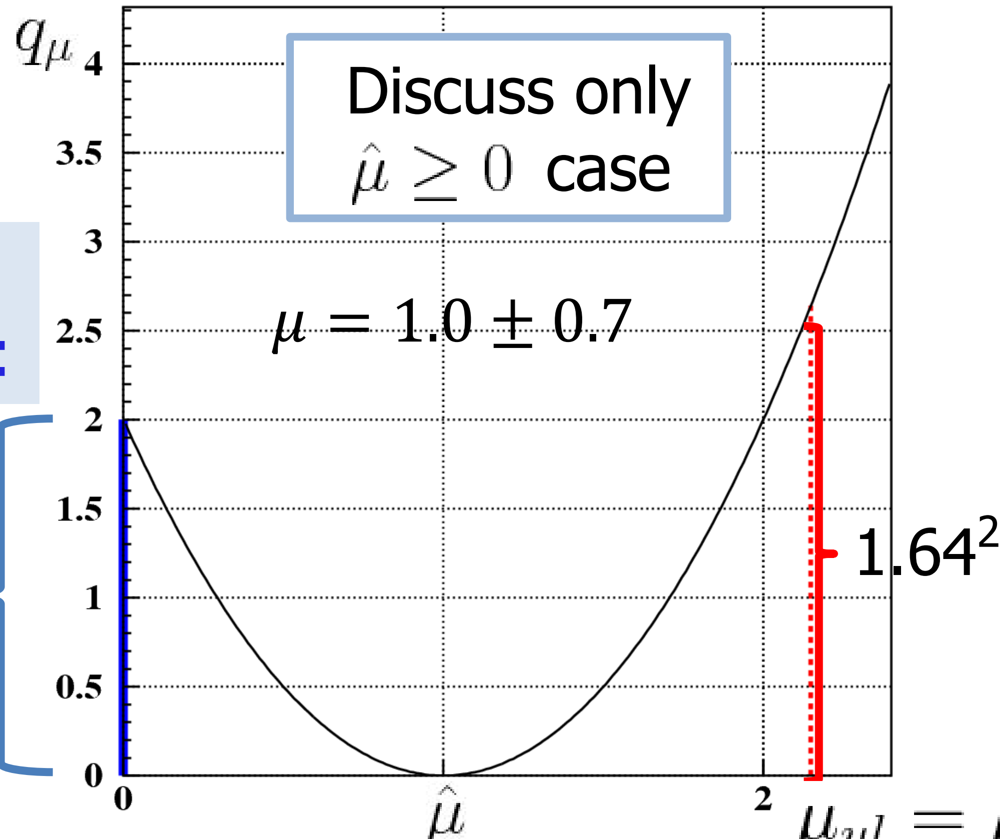
Upper limit $q_\mu$ (set to 0 for $\hat{\mu} \geq \mu$)

# Asymptotics: **CCGV paper arXiv:1007.1727, based on results of Wilks and Wald**

Large sample approximation:

$$q_\mu = -2\ln\lambda(\mu) \approx \frac{(\mu - \hat{\mu})^2}{\sigma^2}$$

$\hat{\mu}$ follows gaussian around true value




For discovery:

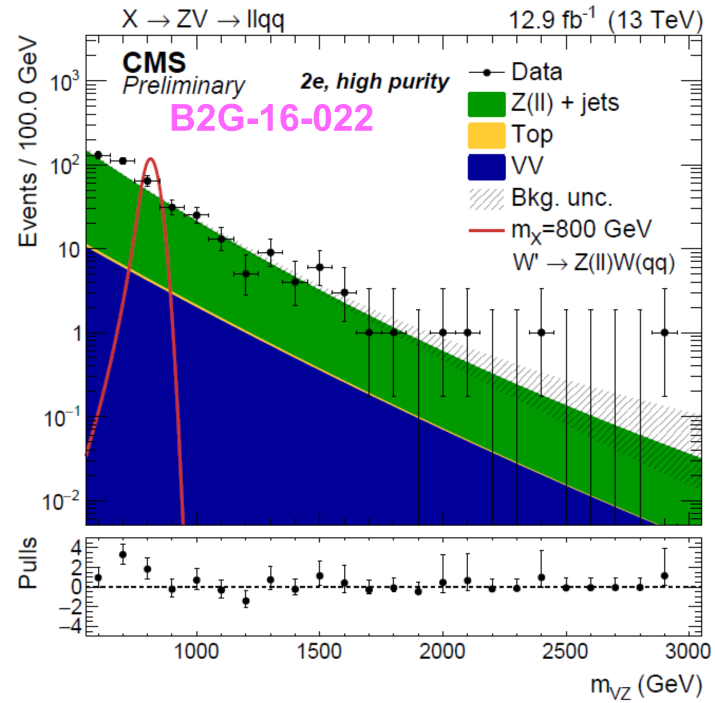
$q_0$

$Z = \sqrt{q_0}$

$\equiv$ Significance

$$\mu_{ul} = \hat{\mu} + 1.64\sigma$$

Upper limit @95% C.L.



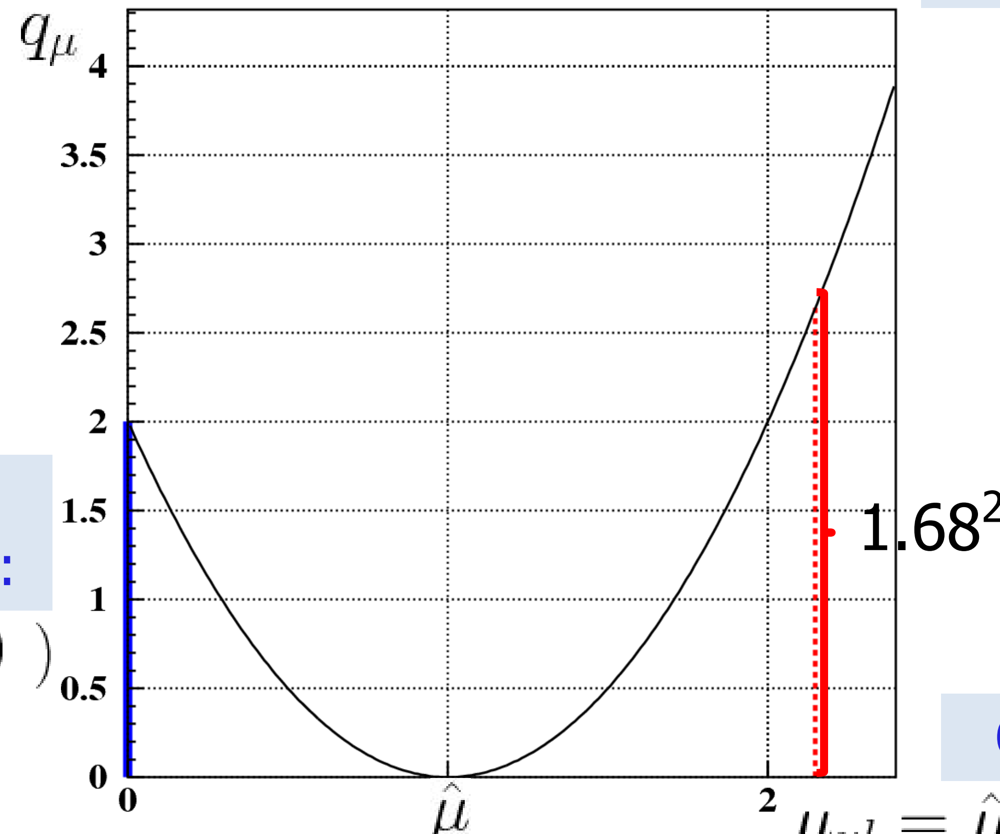
# CLs criterion

For our example with $\hat{\mu} \geq 0$ and asymptotics

$\hat{\mu}_r$ := best fit value for repeated experiments

Find $\mu_{ul}$ such that $\dfrac{p(\hat{\mu}_r \leq \hat{\mu} | \mu_{ul})}{p(\hat{\mu}_r \leq \hat{\mu} | \, 0 \, )} \leq 5\%$

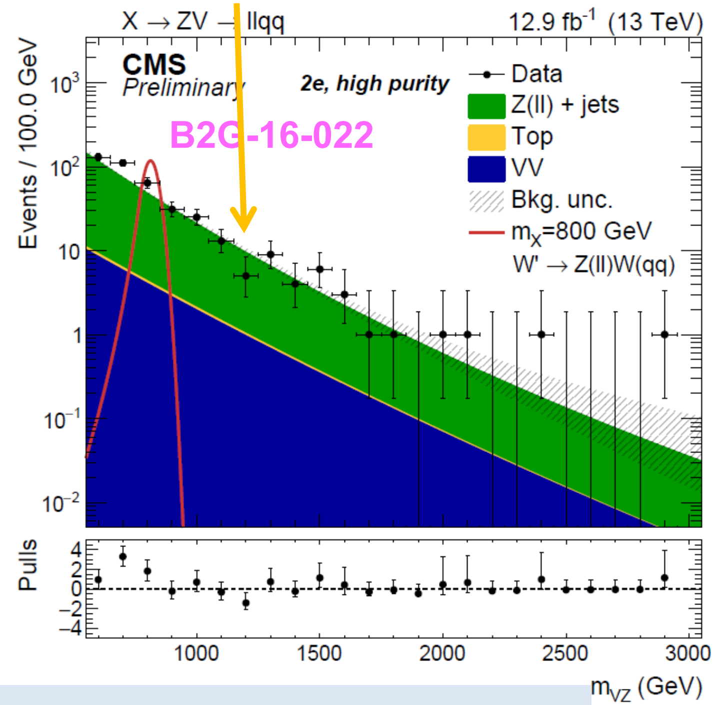
To have less strong limits in case of ↘ fluctuations




For this experiment:

$p(\hat{\mu}_r \leq \hat{\mu} | \, 0 \, )$
=92%

$\mu_{ul} = \hat{\mu} + 1.68\sigma$




CLs Upper limit @95% C.L.

**We use** Likelihood ratio **→ could apply Feldman Cousins procedures to have proper frequentist handling also for $\hat{\mu}$ < 0 → no need for overcovering CLs**

# CLs criterion – description in CMS papers

- 'CLs' should be called a "criterion" or a "prescription" but not a method because otherwise one gets the false impression that much older basic concepts like p-values were only invented with CLs.

- Papers to cite:
  - A.L. Read, "Presentation of search results: the CLs technique", J. Phys. G. 28 (2008) 2693.
  - T. Junk, "Confidence level computation for combining searches with small statistics", NIM A 434 (1999) 435.
  - CMS and ATLAS Collaborations, "Procedure for the LHC Higgs boson search combination in Summer 2011", https://cds.cern.ch/record/1379837, CMS-NOTE-2011-005, ATL-PHYS-PUB-2011-11, CERN, 2011.
  - CCGV: G. Cowan, K. Cranmer, E. Gross, and O. Vitells, "Asymptotic formulae for likelihood-based tests of new physics", EPJC 71 (2011) 1554.

foundational

If toys are used

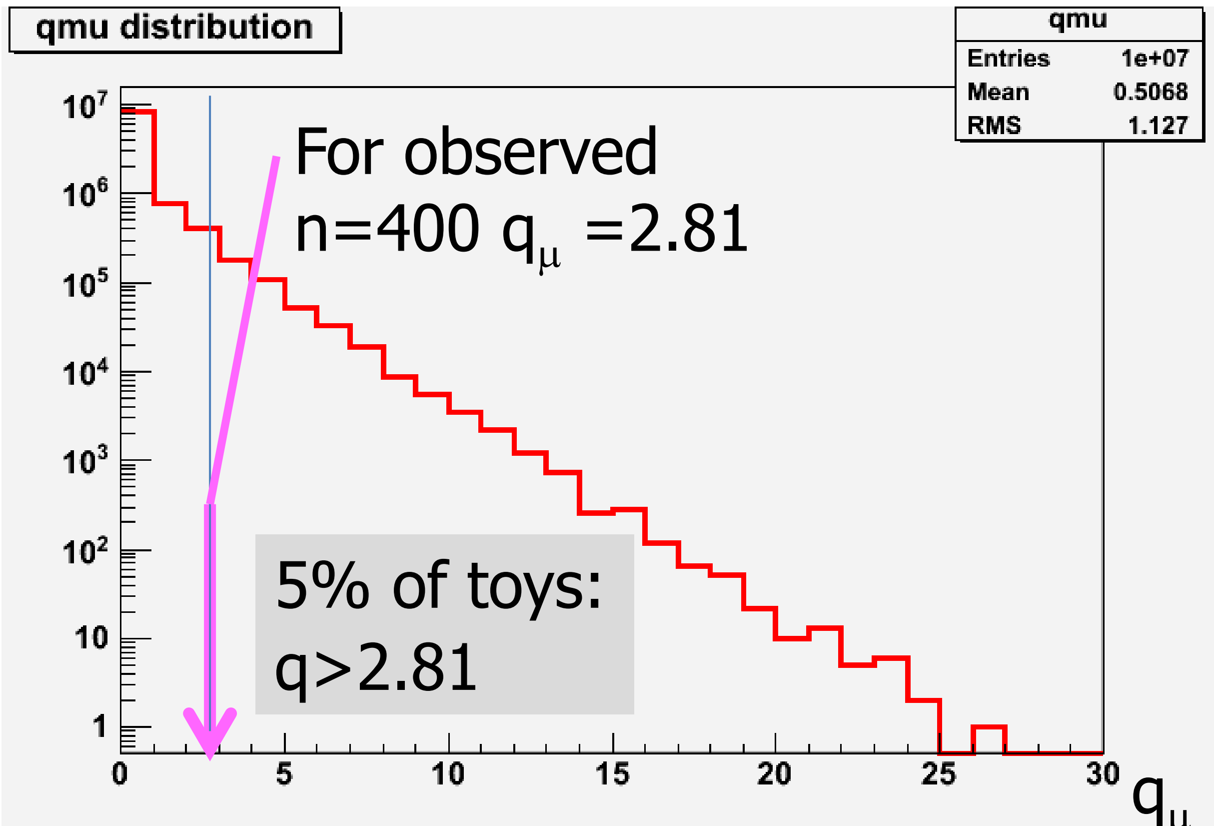
Definition of LHC test statistics + asymptotic formulae

- Please specify for each nuisance parameter assumed uncertainty effect on expected event rates: log-normal, normal, etc.

# Limit for single event count with LHC test statistics

$$L(\mu) = e^{-(\mu+b)} \cdot (\mu+b)^n/n! \qquad q_\mu = \begin{cases} -2\ln\left(\frac{L(\mu)}{L(\hat{\mu})}\right) & \hat{\mu} \leq \mu \\ 0 & \hat{\mu} > \mu \end{cases}$$

- Example: b=400.; n=400
- $q_\mu$ distribution for this b and $\mu$ = 34.5 from toys:




**→Toys:**

$\mu_{limit}$ = 34.5

**→Asymptotics:**

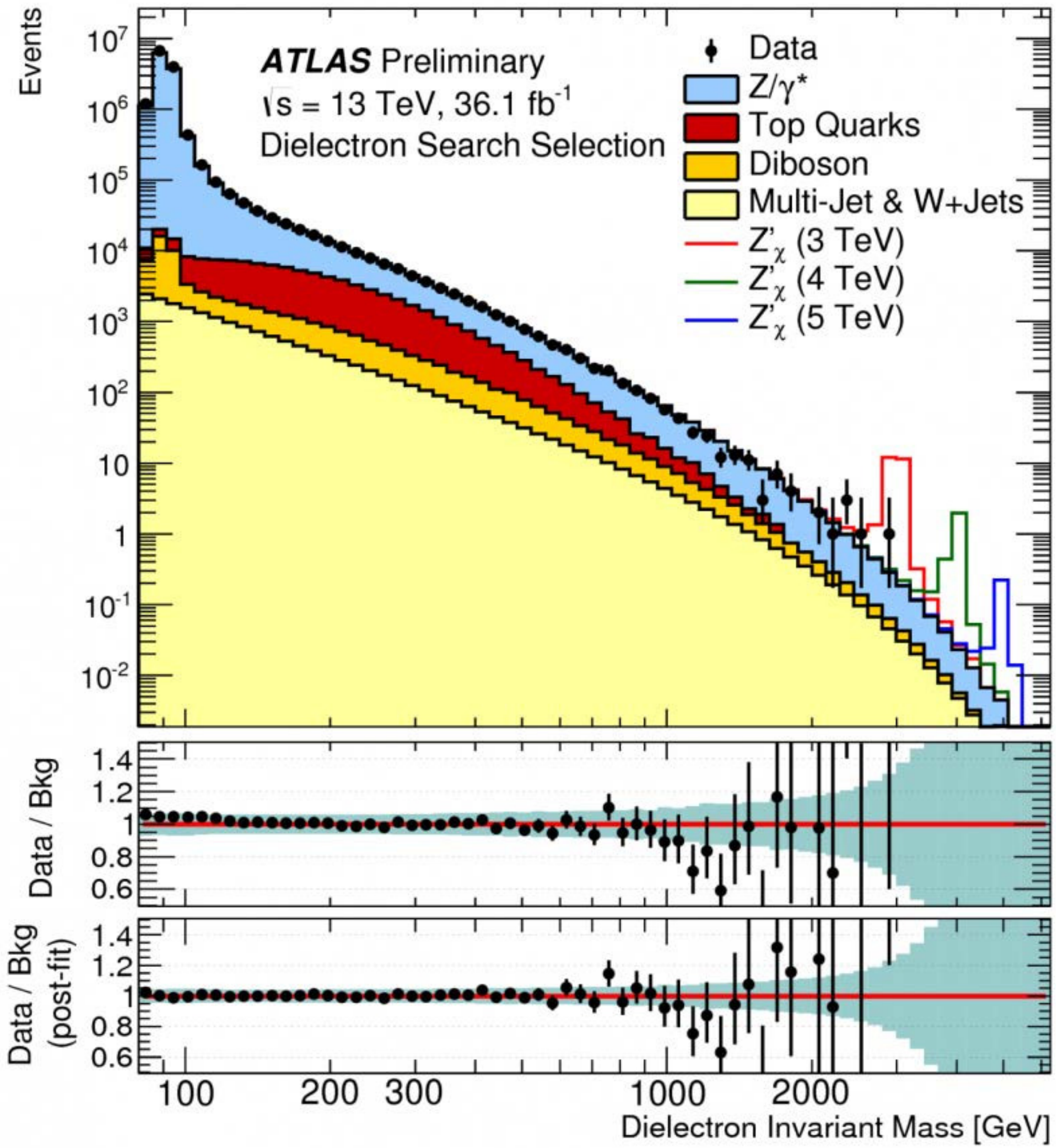
$q_{\mu\ |n=400} = 1.64^2 = 2.69$

Vary $\mu$ until eq. fulfilled

→ $\mu_{limit}$ = 33.7

**→verify asymptotics with toys!**

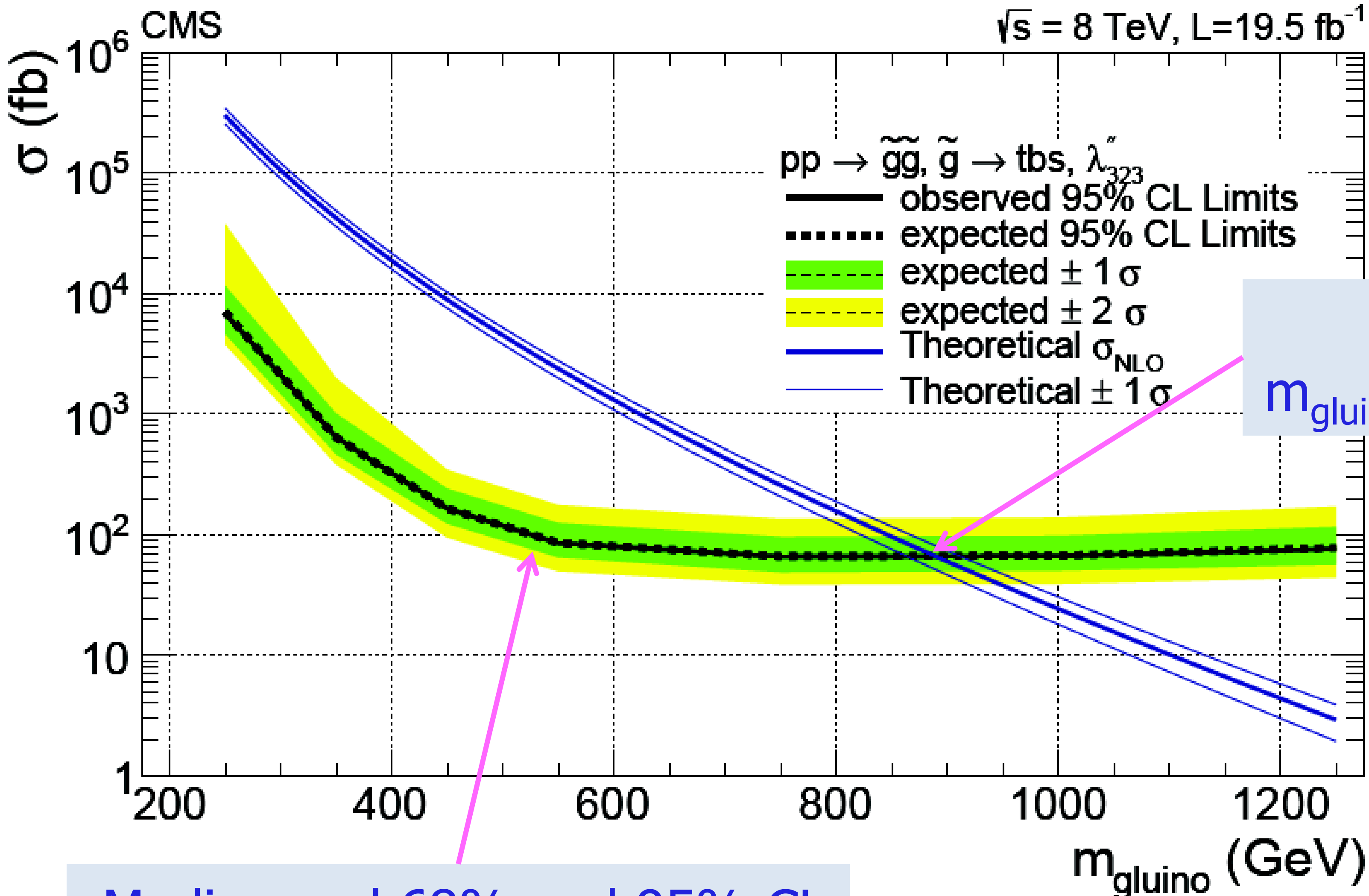
# Typical LHC limits → Scan vs mass of new particles




ATLAS-CONF-2016-045

# Typical LHC limit plots: 1D

CMS SUS-13-013

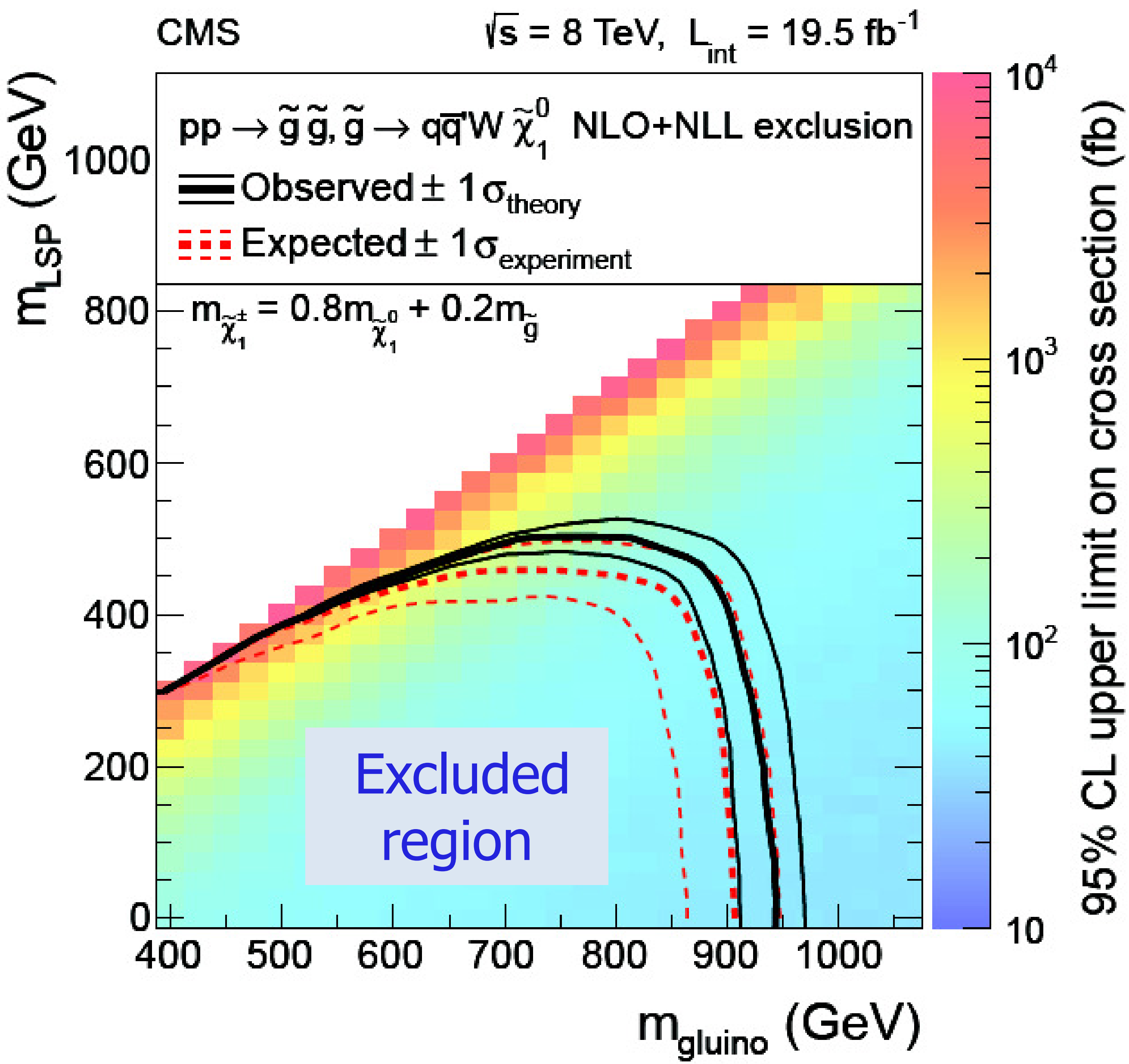



Exclude $m_{gluino} \leq 880$ GeV

Median and 68% and 95% CL expected bands from toys

# Typical LHC limit plots: 2D

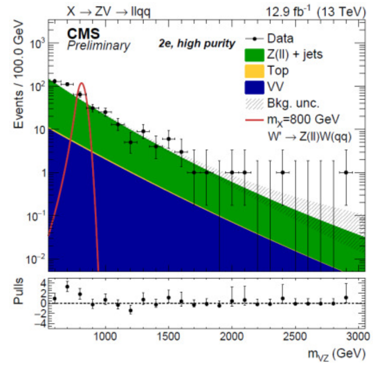
CMS SUS-13-013



# ANALYSERS tasks for typical LHC limit setting

Nuisance pars

$$L(\mu,\boldsymbol{\theta}) = \prod_{j=1}^{N} \frac{(\mu s_j + b_j)^{n_j}}{n_j!} e^{-(\mu s_j + b_j)} \cdot \text{L(control data)} \cdot \text{PDFs}(\theta)$$




CMS B2G-16-022

- Systematics:
  - define set of independent nuisance parameters θ
  - with proper PDFs and mapping to expected rates: $s_j\,(\theta)$, $b_j\,(\theta)$
    Linearization? Template morphing? (understand what is done!)
  - Cross section limits: multiplicative factors (luminosity, eff., etc.)
    → ≈ gaussian uncertainty for ln(x-sec) →"**Log-normal uncertainty**"

- Check everything!
  - GOF-tests, shifts of nuisance parameters (best fit values or profiled), verify asymptotic limits with toys, etc.

# Bayesian upper limit estimation in a 

Nuisance pars

Likelihood

$$L(\mu, \boldsymbol{\theta}) = \prod_{j=1}^{N} \frac{(\mu s_j + b_j)^{n_j}}{n_j!} e^{-(\mu s_j + b_j)} \cdot \text{L(control data)}$$

Priors

**Posterior density:**

$$p(\mu, \boldsymbol{\theta}) = \frac{L(\mu, \boldsymbol{\theta}) \cdot \pi(\mu) \cdot \pi(\boldsymbol{\theta})}{\int L(\mu, \boldsymbol{\theta}) \cdot \pi(\mu) \cdot \pi(\boldsymbol{\theta}) \, d\mu \, d\boldsymbol{\theta}}$$

**Marginalise = integrate out θ**

$$p_m(\mu) = \int p(\mu, \boldsymbol{\theta}) \, d\boldsymbol{\theta}$$

∫ Usually done with Markov-Chain MC

**Limit determination:**

$$CL = P(\mu \leq \mu_{limit}) = 1 - \alpha = \int_0^{\mu_{limit}} p_m(\mu) d\mu$$

**Priors:** $\pi(\mu)=1$ popular at LHC; ⚠ priors change under parameter trafo; → study result sensitivity to prior choices

**Software:** BAT toolkit A. Caldwell et. al.; theta (J. Ott)

# Bayesian upper limit estimation in a 

Nuisance pars

Likelihood

$$L(\mu,\boldsymbol{\theta}) = \prod_{j=1}^{N} \frac{(\mu s_j + b_j)^{n_j}}{n_j!} e^{-(\mu s_j + b_j)} \cdot \text{L(control data)}$$

Priors

Posterior density:

$$p(\mu,\boldsymbol{\theta}) = \frac{L(\mu,\boldsymbol{\theta}) \cdot \pi(\mu) \cdot \pi(\boldsymbol{\theta})}{\int L(\mu,\boldsymbol{\theta}) \cdot \pi(\mu) \cdot \pi(\boldsymbol{\theta})\, d\mu\, d\boldsymbol{\theta}}$$

Marg … integ … th … C

dete … $d\mu$

Priors … ces

- In LHC practice Bayesian and frequentist upper limits seem to agree often fairly well → "are in asymptotic Nirvana" (Bob Cousins)
- If they not agree (usually for very low stat.) they probably "address different questions" (Bob Cousins)
- If you use Bayesian estimation you should run for some exemplary assumed true $\mu$ values toy experiments to check the coverage!

Software: BAT toolkit A. Caldwell et. al.); theta (J. Ott)

# Interval estimation summary I:

- **Interval estimation @LHC times** is relying on sophisticated software → good to remember foundations:
  - Frequentist: consider all outcomes x for any true $\mu$ → $[\mu_l,\mu_h]_x$
  - Bayesian: consider only x of current exp. → posterior $\mu$ density

- **Standard** ATLAS-CMS 95% CL upper limit procedure based on:
  - Likelihood ratio ("LHC style") with profiling all syst. uncertainties
  - Applying at the end CLs prescription (political agreement)
  - My personal impression: for most LHC analyses Bayesian treatment with flat priors gives very similar results

# Interval estimation summary II:

- Most important tasks for analysers:
  - Set up properly analysis and perform checks throughout (e.g. GOF tests in all control regions)
  - Not discussed today: optimizing the sensitivity (with MCs)
    Note: discovery and best limit need different optimisation!

## Final recommendations:

→ Participate in RooFit/RooStats tutorial (L. Moneta)
→ Statistics & Systematic treatments:
  → read other analysis papers (from same/other experiment)
  → present your analysis often in working group meetings!

# Backup slides

# Calculation of Poisson sums

- Sum over Poisson terms is related to $\chi^2$ distribution with number-of-degrees of freedom "k":

$$\chi^2(x;k)=\frac{x^{k/2-1}e^{-x/2}}{2^{k/2}\Gamma(k/2)} \qquad P(N;\mu)=\frac{e^{-\mu}\mu^N}{N!}$$

- Poisson sum equals integral over $\chi^2$ distribution (partial integration)

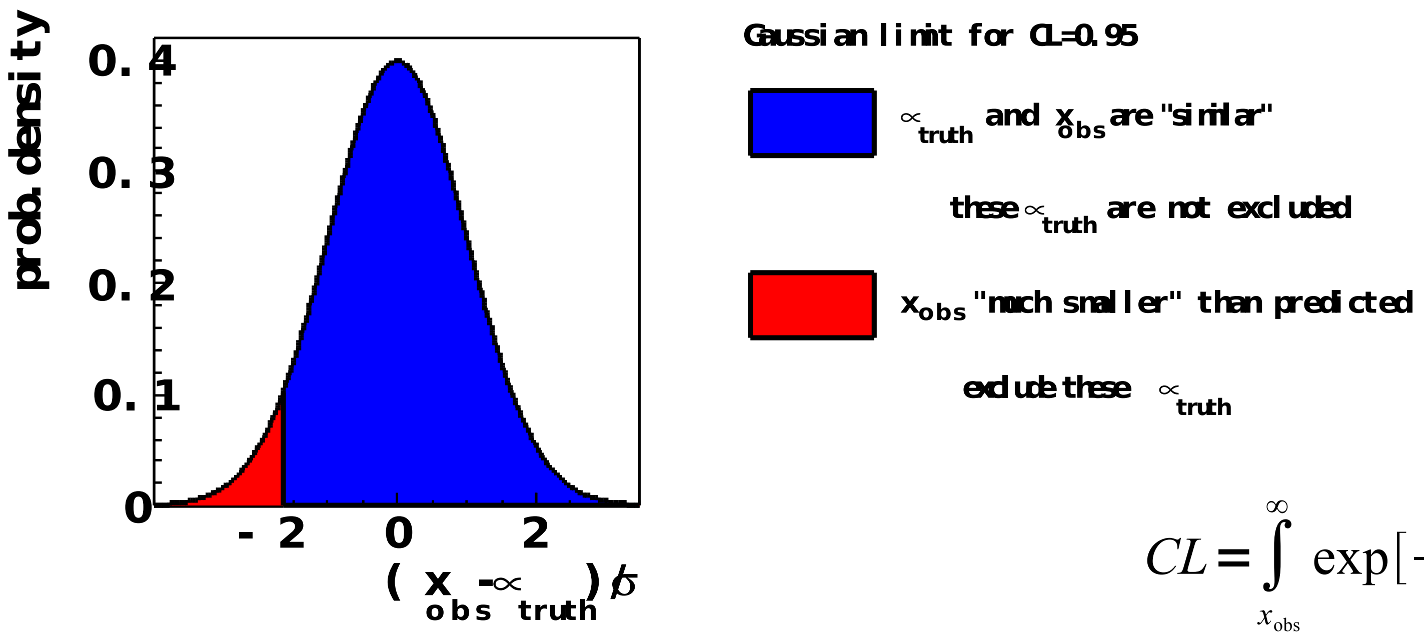
$$\alpha(\mu,N)=\int_{2\mu}^{\infty}\chi^2(x;2(N+1))\,\mathrm{d}x=\sum_{i=0}^{N}P(i;\mu)$$

- Standard functions for $\chi^2$ integrals:

`α(μ,N)=TMath::Prob(2*μ,2*(N+1))` and

`μ=0.5*TMath::ChisquareQuantile(1-α,2*(N+1))`

# Frequentist upper limit, Gaussian case




$$CL = \int_{x_{obs}}^{\infty} \exp[-1/2(\frac{x-\mu_{truth}}{\sigma})^2]\frac{1}{\sqrt{2\pi}\sigma}\,dx$$

- Fixed σ, measurement $x_{obs}$, parameter of interest $\mu_{truth}$
- Define 95% probability area under Gaussian
- If $\mu_{truth}$ is too large, it is outside the 95% → excluded

# Limits with background, comparison

- Frequentist limit may become “unphysicsal” or “too good”
- Expected limit: sensitivity of the experiment
- $CL_s$ method: normalize to “standard model”, never exclude zero signal
- Disadvantage of $CL_s$? Study coverage