Search and Discovery Statistics in HEP
LECTURE 2

Eilam Gross, Weizmann Institute of Science

This presentation would have not been possible without the tremendous help of the following people throughout many years

Louis Lyons, Alex Read, Glen Cowan ,Kyle Cranmer
Ofer Vitells & Bob Cousins

Confidence Interval and Confidence Level (CL)

CL & CI - Wikipedia

$$\mu = 1.1 \pm 0.3$$

$$\mu = [0.8, 1.4] \ @ \ 68\% \ CL$$

CI=[0.8,1.4]

what does it mean?

- A **confidence interval** (**CI**) is a particular kind of interval estimate of a population parameter.
- Instead of estimating the parameter by a single value, an interval likely to include the parameter is given.
- How likely the interval is to contain the parameter is determined by the **confidence level** or confidence coefficient.
- Increasing the desired confidence level will widen the confidence interval.

Confidence Interval & Coverage

- Say you have a measurement μ_{meas} of μ with μ_{true} being the unknown true value of μ
- Assume you know the probability distribution function $p(\mu_{meas}|\mu)$
- Given the measurement you deduce somehow (based on your statistical model) that there is a 95% Confidence interval $[\mu_1,\mu_2]$.

(it is 95% likely that the μ_{true} is in the quoted interval)

The correct statement:

- In an ensemble of experiments 95% of the obtained confidence intervals will contain the true value of μ.

Upper limit

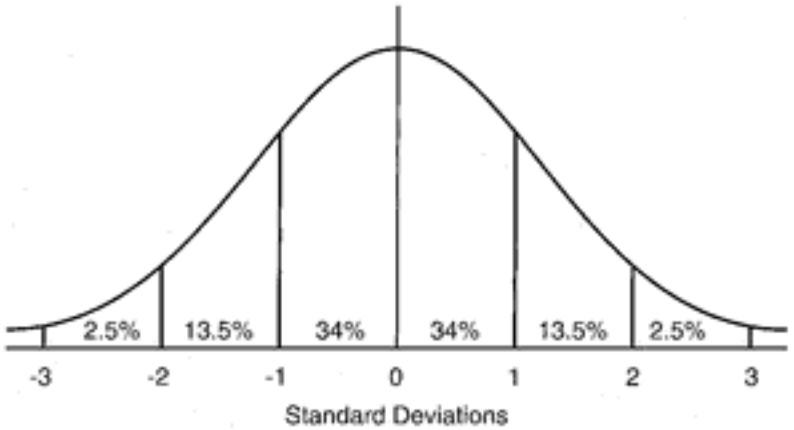
- Given the measurement you deduce somehow (based on your statistical model) that there is a 95% Confidence interval $[0,\mu_{up}]$.
- This means: our interval contains $\mu=0$ (no Higgs)
- We therefore deduce that $\mu<\mu_{up}$ at the 95% Confidence Level (CL)
- μ_{up} is therefore an upper limit on μ
- If $\mu_{up}<1 \rightarrow$
  $\sigma(m_H)<\sigma_{SM}(m_H) \rightarrow$
  a SM Higgs with a mass m_H is excluded at the 95% CL

Confidence Interval & Coverage

- Confidence Level: A CL of (e.g.) 95% means that in an ensemble of experiments, each producing a confidence interval, 95% of the confidence intervals will contain the true value of μ
- Normally, we make one experiment and try to estimate from this one experiment the confidence interval at a specified CL
- If in an ensemble of (MC) experiments our estimated Confidence Interval fail to contain the true value of μ 95% of the cases (for every possible μ) we claim that our method undercover
- If in an ensemble of (MC) experiments our estimated Confidence Interval contains the true value of μ more than 95% of the cases (for every possible μ) we claim that our method overcover (being conservative)
- If in an ensemble of (MC) experiments the true value of μ is covered within the estimated confidence interval , we claim a coverage

How to deduce a CI?

- One can show that if the data is distributed normal around the average i.e. P(data | μ)=normal



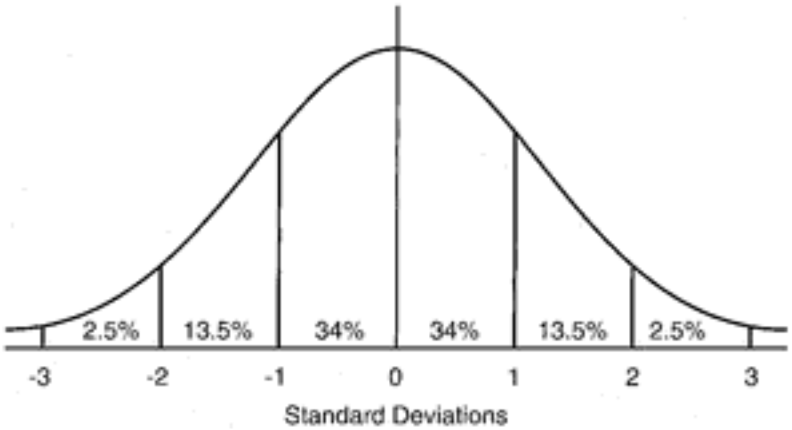
$$f(x \mid \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

then one can construct a 68% CI around the estimator of μ to be

$$\hat{x} \pm \sigma$$

However, not all distributions are normal, many distributions are even unknown and coverage might be a real issue

How to deduce a CI?



- One can show that if the data is distributed normal around the average i.e. P(data | μ)=normal

$$f(x \mid \mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

then one can construct a 68% CI around the estimator of μ to be

$$\hat{x} \pm \sigma$$

However, not all distributions are normal, many distributions are even unknown and coverage might be a real issue

- One may construct many 68% intervals…. $CI = [\mu_L, \mu_U]$

$$\int_{\mu_L}^{\mu_U} f(x \mid \hat{x})dx = 68\%$$

- Which one has a full coverage?
- How can we guarantee a coverage

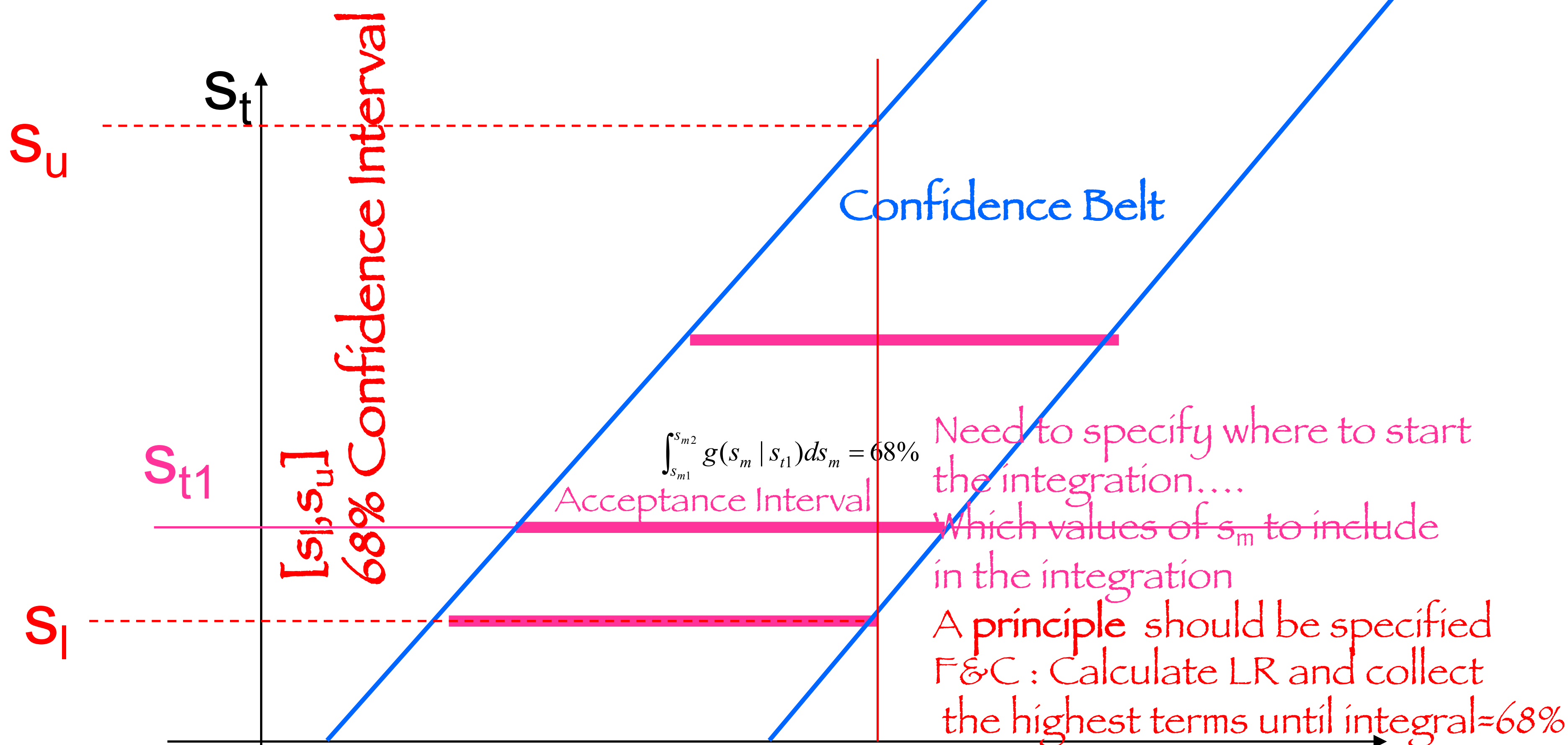
- The QUESTION is NOT how to construct a CI, it is
- HOW TO CONSTRUCT A CI WHICH HAS A COVERAGE @ THE 68% CL

The Frequentist Game a 'la Feldman & Cousins

Or

How to ensure a Coverage

Neyman Construction



$[s_l,s_u]$ 68% Confidence Interval
In **68%** of the experiments the derived **C.I. contains** the **unknown true value of s**

- With Neyman Construction we guarantee a coverage via construction, i.e. for any value of the unknown true s, the Construction Confidence Interval will cover s with the correct rate.

The Flip Flop Way of an Experiment

- The most intuitive way to analyze the results of an experiment would be
  - Construct a test statistics
    e.g. $Q(x) \sim L(x|H_1) / L(x|H_0)$
  - If the significance of the measured $Q(x_{obs})$, is less than 3 sigma, derive an upper limit (just looking at tables), if the result is >5 sigma (and some minimum number of events is observed….), derive a discovery central confidence interval for the measured parameter (cross section, mass….) …..
- **This Flip Flopping policy leads to undercoverage:**
  Is that really a problem for Physicists?
  Some physicists say, for each experiment quote always two results, an upper limit, and a (central?) discovery confidence interval
- Many LHC analyses report both ways.

Frequentist Paradise – F&C Unified with Full Coverage

- Frequentist Paradise is certainly made up of an interpretation by constructing a confidence interval in brute force ensuring a coverage!
- This is the Neyman confidence interval adopted by F&C….

- The motivation:
  - Ensures Coverage
  - Avoid Flip-Flopping – an ordering rule determines the nature of the interval (1-sided or 2-sided depending on your observed data)
  - Ensures Physical Intervals

- Let the test statistics be

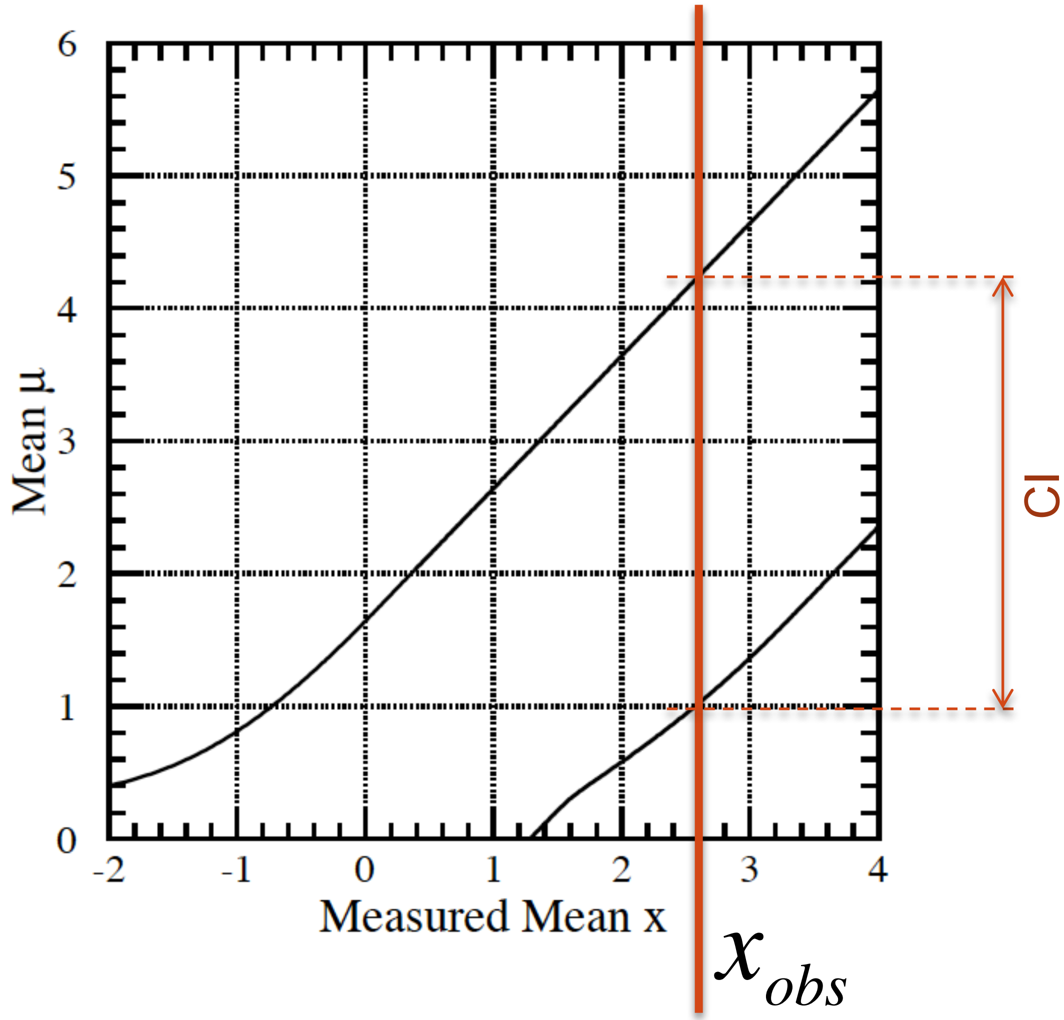
$$Q = \begin{cases} \dfrac{L(s+b)}{L(\hat{s}+b)} & \hat{s} > 0 \\ \dfrac{L(s+b)}{L(b)} & \hat{s} \leq 0 \end{cases}$$

  where $\hat{s}$ is the **physically allowed** mean s that maximizes $L(\hat{s}+b)$
  (If $\hat{s}<0$, denominator is set to $L(b)$.)

- Order by taking the 68% highest Qs

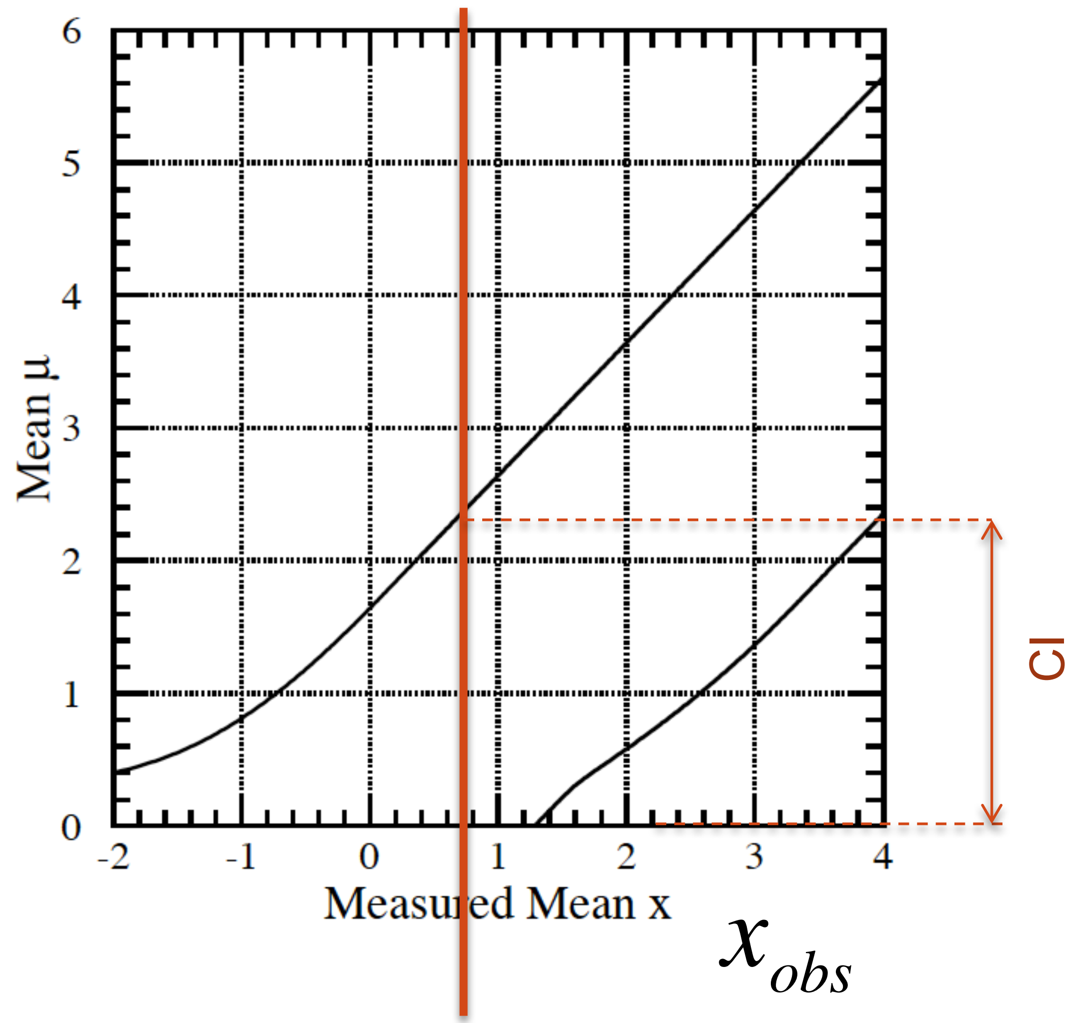
How to tell an Upper limit from a Measurement without Flip Flopping

- A measurement (2 sided)



How to tell an Upper limit from a Measurement without Flip Flopping

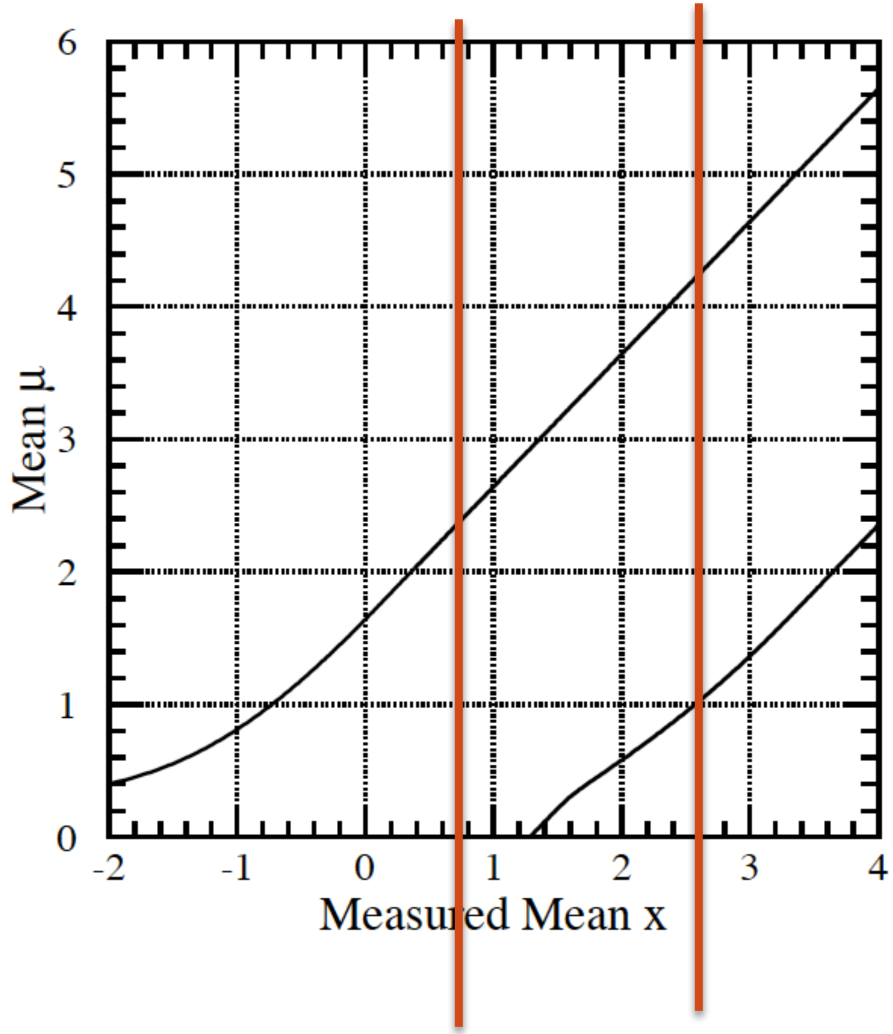
- An upper limit (1 sided)



How to tell an Upper limit from a Measurement without Flip Flopping

- Your observed result will tell you if it's a measurement or an upper limit

- But how to deal with systematics?



Nuisance Parameters

or Systematics

Nuisance Parameters (Systematics)

- There are two kinds of parameters:
  - Parameters of interest (signal strength… cross section… μ)
  - Nuisance parameters (background (b), signal efficiency, resolution, energy scale,…)
- The nuisance parameters carry systematic uncertainties
- There are two related issues:
  - Classifying and estimating the systematic uncertainties
  - Implementing them in the analysis
- The physicist must make the difference between cross checks and identifying the sources of the systematic uncertainty.
  - Shifting cuts around and measure the effect on the observable… Very often the observed variation is dominated by the statistical uncertainty in the measurement.

Implementation of Nuisance Parameters

- Implement by marginalizing (Bayesian) or profiling (Frequentist)
  - One can also use a frequentist test statistics (PL) while treating the NPs via marginalization (Hybrid, Cousins & Highland way)
- Marginalization (Integrating))
  - Integrate the Likelihood, L, over possible values of nuisance parameters (weighted by their prior belief functions -- Gaussian,gamma, others...)
  - Consistent Bayesian interpretation of uncertainty on nuisance parameters

Integrating Out The Nuisance Parameters (Marginalization)

- Our degree of belief in μ is the sum of our degree of belief in μ given θ(nuisance parameter), over "all" possible values of θ
- That's a Bayesian way

$$\boxed{p(\mu\,|\,x) = \int p(x\,|\,\mu,\theta)\pi(\theta)\,d\theta = \int L(\mu,\theta)\pi(\theta)\,d\theta}$$

Credible Interval $CI = [0, \mu_{up}]$

$$\boxed{0.95 = \int_0^{\mu_{up}} p(\mu\,|\,x)\,d\mu}$$

Nuisance Parameters (Systematisc)

- Neyman Pearson (NP) Likelihood Ratio:

$$q^{NP} = -2\ln\frac{L(b(\theta))}{L(s+b(\theta))}$$

- Either Integrate the Nuisance parameters

prior

$$q^{NP}_{Hybrid} = \frac{\int L\big(s+b(\theta)\big)\pi(\theta)\,d\theta}{\int L\big(b(\theta)\big)\pi(\theta)\,d\theta}$$

Cousins & Highland

- Or profile them

$$q^{NP} = -2\ln\frac{L\left(b(\hat{\hat{\theta}}_0)\right)}{L\left(s+b(\hat{\hat{\theta}}_1)\right)}$$

$$\hat{\hat{\theta}}_0 = MLE_{\mu=0}\ of\ L(b(\theta))$$

$$\hat{\hat{\theta}}_1 = MLE_{\mu=1}\ of\ L(s+b(\theta))$$

Nuisance Parameters and Subsidiary Measurements

- Usually the nuisance parameters are auxiliary parameters and their values are constrained by auxiliary measurements
- Example

$$n \sim \mu s(m_H) + b \qquad \langle n \rangle = \mu s + b$$

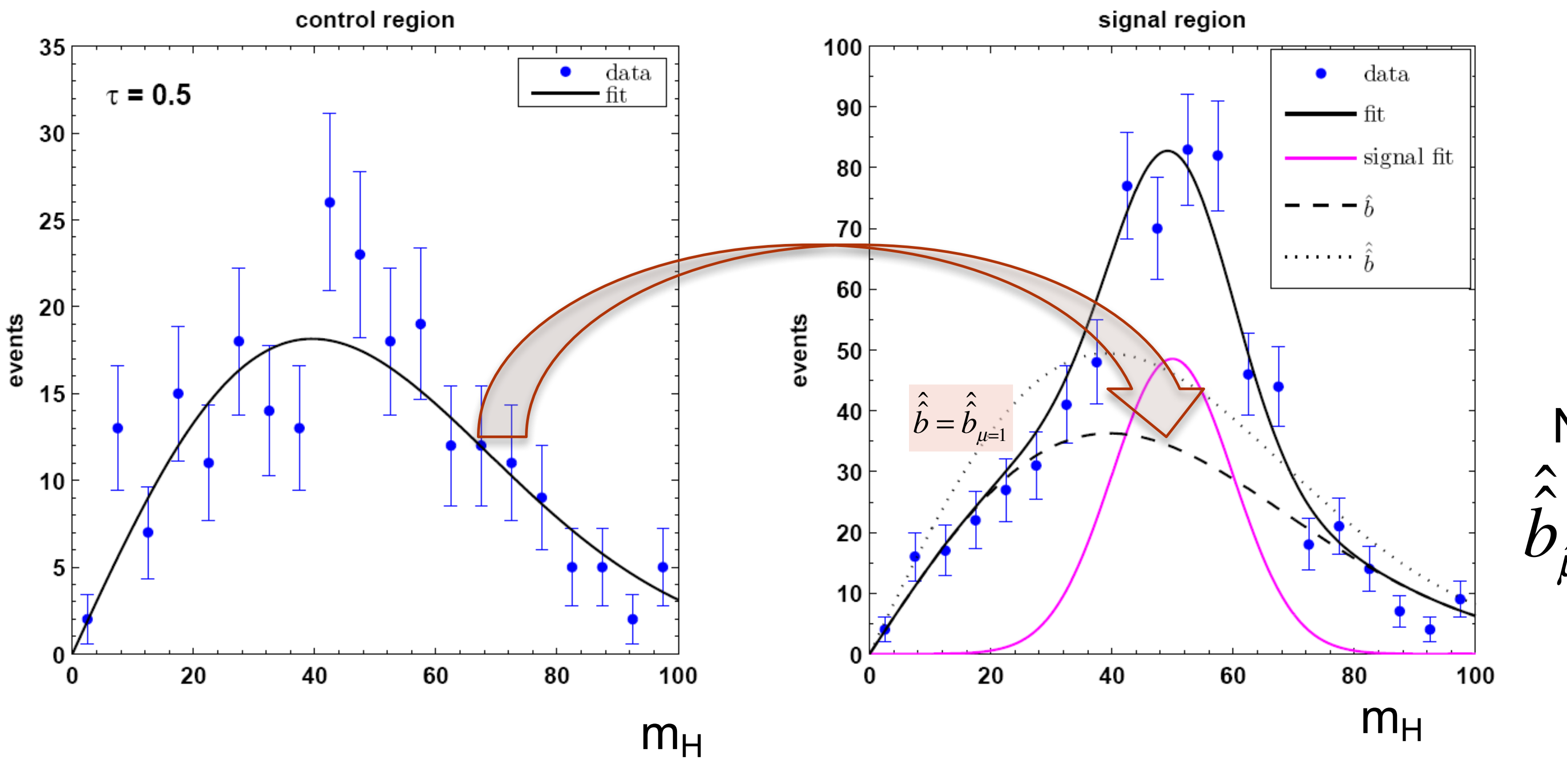
$$m = \tau b$$

$$\boxed{L\big(\mu \cdot s + b(\theta)\big) = Poisson\big(n; \mu \cdot s + b(\theta)\big) \cdot Poisson\big(m; \tau b(\theta)\big)}$$

Mass shape as a discriminator

$$n:\ \mu s(m_H)+b \qquad m \sim \tau b$$

$$L\big(\mu\cdot s+b(\theta)\big)=\prod_{i=1}^{nbins} Poisson\big(n_i;\mu\cdot s_i+b_i(\theta)\big)\cdot Poisson\big(m_i;\tau b_i(\theta)\big)$$



Note

$$\hat{\hat{b}}_{\hat{\mu}}=\hat{b}$$

Profile Likelihood with Nuisance Parameters

$$q_\mu = -2ln\frac{L(\mu s+\hat{\hat{b}}_\mu)}{L(\hat{\mu}s+\hat{b})}$$

$$q_\mu = -2ln\frac{\max_b L(\mu s+b)}{\max_{\mu,b} L(\mu s+b)}$$

$$q_\mu = q_\mu(\hat{\mu}) = -2ln\frac{L(\mu s+\hat{\hat{b}}_\mu)}{L(\hat{\mu}s+\hat{b})}$$

$\hat{\mu}$ MLE of μ

$\hat{b}$ MLE of b

$\hat{\hat{b}}_\mu$ MLE of b fixing μ

$\hat{\hat{\theta}}_\mu$ MLE of θ fixing μ

Wilks theorem in the presence of NPs

- Given n parameters of interest and any number of NPs, then

$$\boxed{\begin{gathered}\lambda(\alpha_i) \equiv \frac{L(\alpha_i, \hat{\hat{\theta}}_j)}{L(\hat{\alpha}_i, \hat{\theta}_j)} \\ q(\alpha_i) \equiv -2\log\lambda(\alpha_i) \sim \chi_n^2\end{gathered}}$$

Classification of Test Statistics

CCGV

Test Statistics	Purpose	Experession	LR
q_0	discovery of positive signal	$q_0 = \begin{cases} -2\ln\lambda(0) & \hat{\mu} \geq 0 \\ 0 & \hat{\mu} < 0 \end{cases}$	$\lambda(0) = \dfrac{L(0,\hat{\hat{\theta}}_0)}{L(\hat{\mu},\hat{\theta})}$
t_μ	2-sided measurement	$t_\mu = -2\ln\lambda(\mu)$	$\lambda(\mu) = \dfrac{L(\mu,\hat{\hat{\theta}}_\mu)}{L(\hat{\mu},\hat{\theta})}$
$\tilde{t}_\mu$	avoid negative signal (FC)	$\tilde{t}_\mu = -2\ln\tilde{\lambda}(\mu)$	$\tilde{\lambda}(\mu) = \begin{cases} \dfrac{L(\mu,\hat{\hat{\theta}}_\mu)}{L(\hat{\mu},\hat{\theta})} & \hat{\mu} \geq 0 \\ \dfrac{L(\mu,\hat{\hat{\theta}}_\mu)}{L(0,\hat{\hat{\theta}}_0)} & \hat{\mu} < 0 \end{cases}$
q_μ	exclusion	$q_\mu = \begin{cases} -2\ln\lambda(\mu) & \hat{\mu} \leq \mu \\ 0 & \hat{\mu} > \mu \end{cases}$	
$\tilde{q}_\mu$	exclusion of positive signal	$\tilde{q}_\mu = \begin{cases} -2\ln\tilde{\lambda}(\mu) & \hat{\mu} \leq \mu \\ 0 & \hat{\mu} > \mu \end{cases}$	

Test Statistics	Purpose	Experession	LR
q_0	discovery of positive signal	$q_0 = \begin{cases} -2\ln\lambda(0) & \hat{\mu} \geq 0 \\ 0 & \hat{\mu} < 0 \end{cases}$	$\lambda(0) = \dfrac{L(0,\hat{\hat{\theta}}_0)}{L(\hat{\mu},\hat{\theta})}$
t_μ	2-sided measurement	$t_\mu = -2\ln\lambda(\mu)$	$\lambda(\mu) = \dfrac{L(\mu,\hat{\hat{\theta}}_\mu)}{L(\hat{\mu},\hat{\theta})}$
$\tilde{t}_\mu$	avoid negative signal (FC)	$\tilde{t}_\mu = -2\ln\tilde{\lambda}(\mu)$	$\tilde{\lambda}(\mu) = \begin{cases} \dfrac{L(\mu,\hat{\hat{\theta}}_\mu)}{L(\hat{\mu},\hat{\theta})} & \hat{\mu} \geq 0 \\ \dfrac{L(\mu,\hat{\hat{\theta}}_\mu)}{L(0,\hat{\hat{\theta}}_0)} & \hat{\mu} < 0 \end{cases}$
q_μ	exclusion	$q_\mu = \begin{cases} -2\ln\lambda(\mu) & \hat{\mu} \leq \mu \\ 0 & \hat{\mu} > \mu \end{cases}$	
$\tilde{q}_\mu$	exclusion of positive signal	$\tilde{q}_\mu = \begin{cases} -2\ln\tilde{\lambda}(\mu) & \hat{\mu} \leq \mu \\ 0 & \hat{\mu} > \mu \end{cases}$	

Downward fluctuations of the BG do not serve as an evidence against the BG

Test Statistics	Purpose	Experession	LR
q_0	discovery of positive signal	$q_0 = \begin{cases} -2\ln\lambda(0) & \hat{\mu} \ge 0 \\ 0 & \hat{\mu} < 0 \end{cases}$	$\lambda(0) = \frac{L(0,\hat{\hat{\theta}}_0)}{L(\hat{\mu},\hat{\theta})}$

Downward fluctuations of the BG do not serve as an evidence against the BG

t_μ	2-sided measurement	$t_\mu = -2\ln\lambda(\mu)$	$\lambda(\mu) = \frac{}{L(\hat{\mu},\hat{\theta})}$
$\tilde{t}_\mu$	avoid negative signal (FC)	$\tilde{t}_\mu = -2\ln\tilde{\lambda}(\mu)$	$\tilde{\lambda}(\mu) = \begin{cases} \frac{L(\mu,\hat{\hat{\theta}}_\mu)}{L(\hat{\mu},\hat{\theta})} & \hat{\mu} \ge 0 \\ \frac{L(\mu,\hat{\hat{\theta}}_\mu)}{L(0,\hat{\hat{\theta}}_0)} & \hat{\mu} < 0 \end{cases}$
q_μ	exclusion	$q_\mu = \begin{cases} -2\ln\lambda(\mu) & \hat{\mu} \le \mu \\ 0 & \hat{\mu} > \mu \end{cases}$	

Upward fluctuations of the signal do not serve as an evidense against the signal

$\tilde{q}_\mu$	exclusion of positive signal	$\tilde{q}_\mu = \begin{cases} & \\ 0 & \hat{\mu} > \mu \end{cases}$	

Asymptotic Distributions

CCGV

$$q_{null}$$

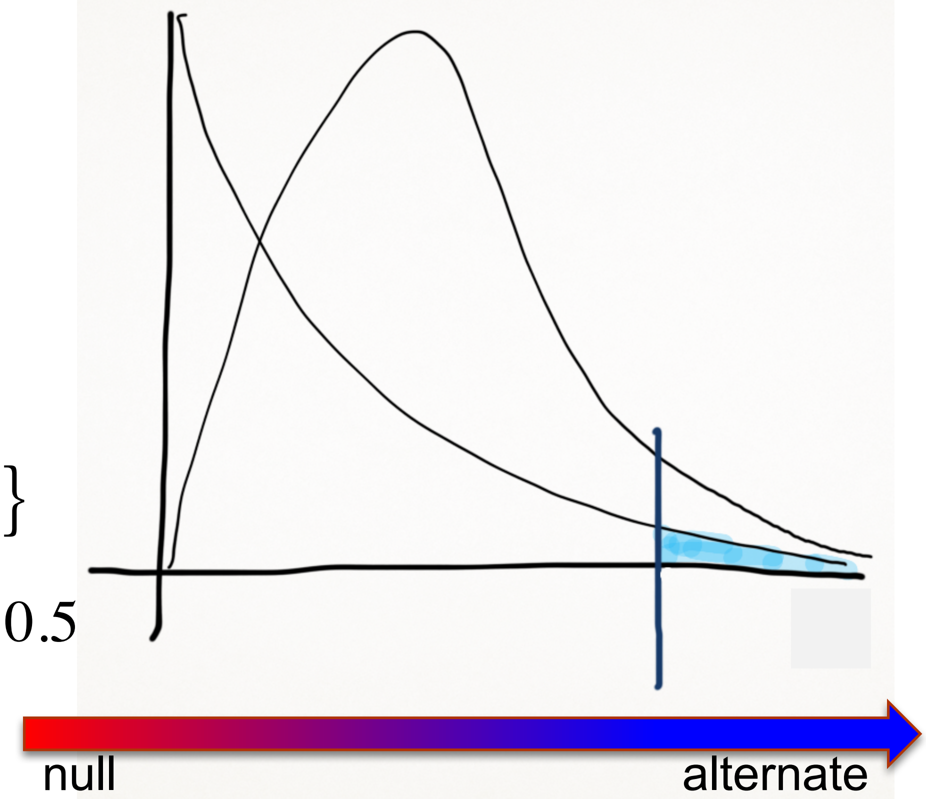
$$f(q_{null} \mid H_{null})$$

$$q_{obs} \equiv q_{null,obs}$$

$$p = \int_{q_{obs}}^{\infty} f(q_{null} \mid H_{null}) dq_{null}$$

$$f(q_{null} \mid H_{alt})$$

$$q_A \equiv q_{null,A} = \begin{array}{l} \{q \mid med\{f(q_{null} \mid H_{alt})\}\} \\ \int_{q_{null,A}}^{\infty} f(q_{null} \mid H_{null}) dq_{null} = 0.5 \end{array}$$



q_{null}

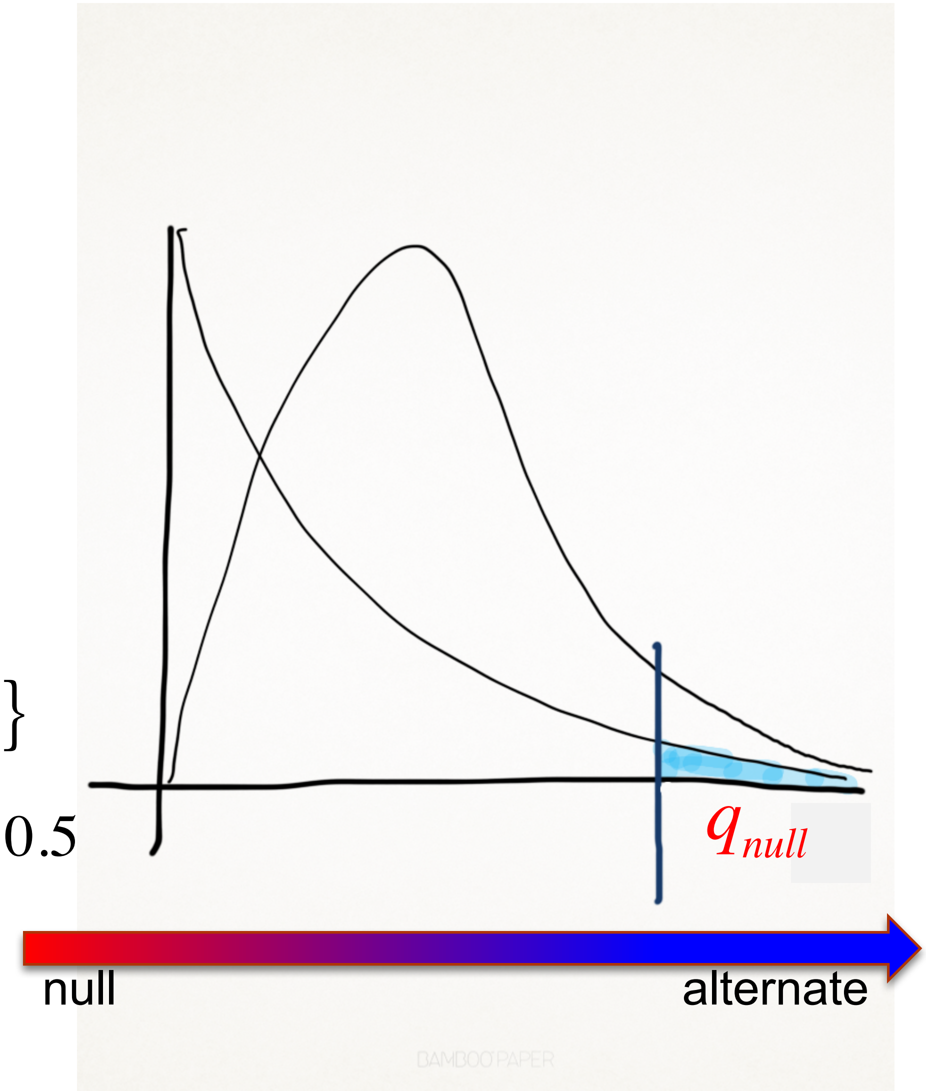
$f(q_{null} | H_{null})$

$q_{obs} \equiv q_{null,obs}$

$$p = \int_{q_{obs}}^{\infty} f(q_{null} | H_{null}) dq_{null}$$

$f(q_{null} | H_{alt})$

$$q_A \equiv q_{null,A} = \begin{array}{l} \{q \,|\, med\{f(q_{null} | H_{alt})\}\} \\ \int_{q_{null,A}}^{\infty} f(q_{null} | H_{null}) dq_{null} = 0.5 \end{array}$$



q_{null}

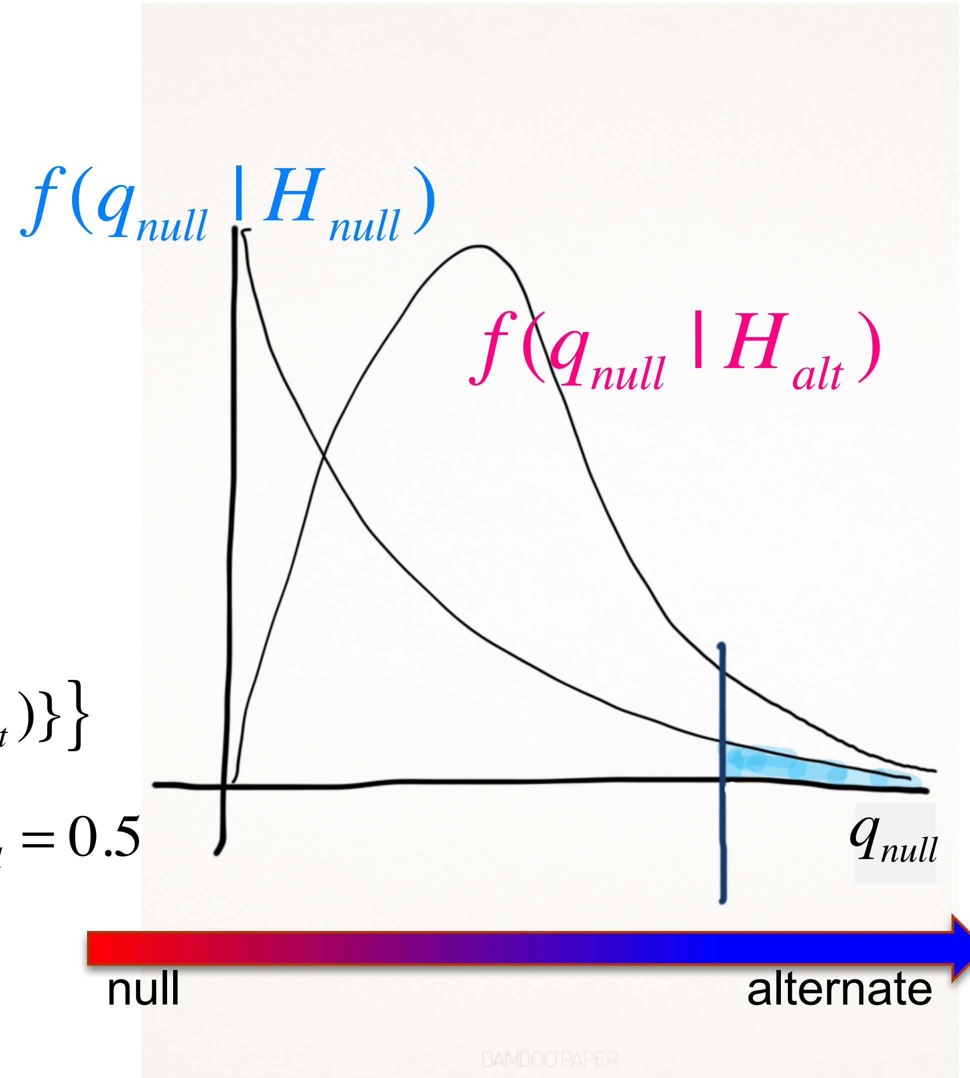
$f(q_{null} | H_{null})$

$q_{obs} \equiv q_{null,obs}$

$$p = \int_{q_{obs}}^{\infty} f(q_{null} | H_{null}) dq_{null}$$

$f(q_{null} | H_{alt})$

$$q_A \equiv q_{null,A} = \begin{array}{c} \{q | med\{f(q_{null} | H_{alt})\}\} \\ \int_{q_{null,A}}^{\infty} f(q_{null} | H_{null}) dq_{null} = 0.5 \end{array}$$



q_{null}

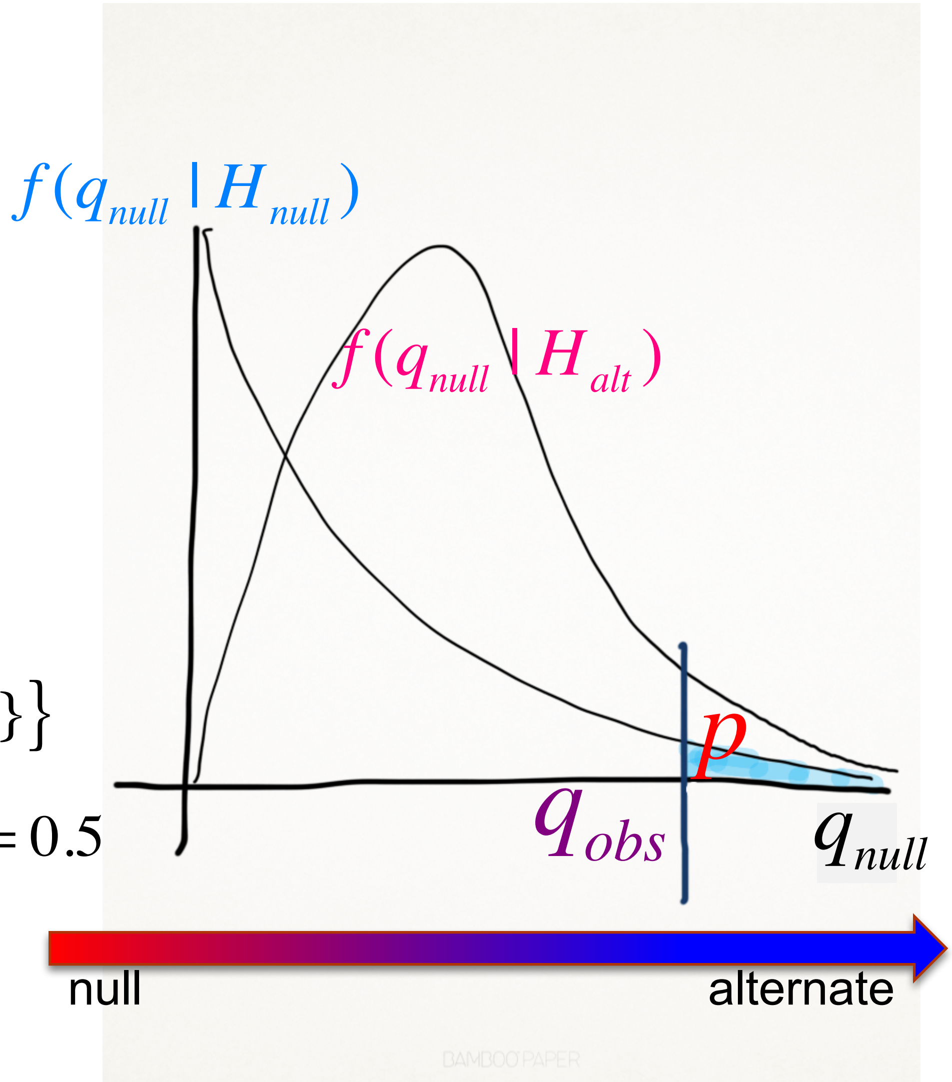
$f(q_{null} \mid H_{null})$

$q_{obs} \equiv q_{null,obs}$

$$p = \int_{q_{obs}}^{\infty} f(q_{null} \mid H_{null}) dq_{null}$$

$f(q_{null} \mid H_{alt})$

$$q_A \equiv q_{null,A} = \begin{array}{c} \{q \mid med\{f(q_{null} \mid H_{alt})\}\} \\ \int_{q_{null,A}}^{\infty} f(q_{null} \mid H_{null}) dq_{null} = 0.5 \end{array}$$



q_{null}

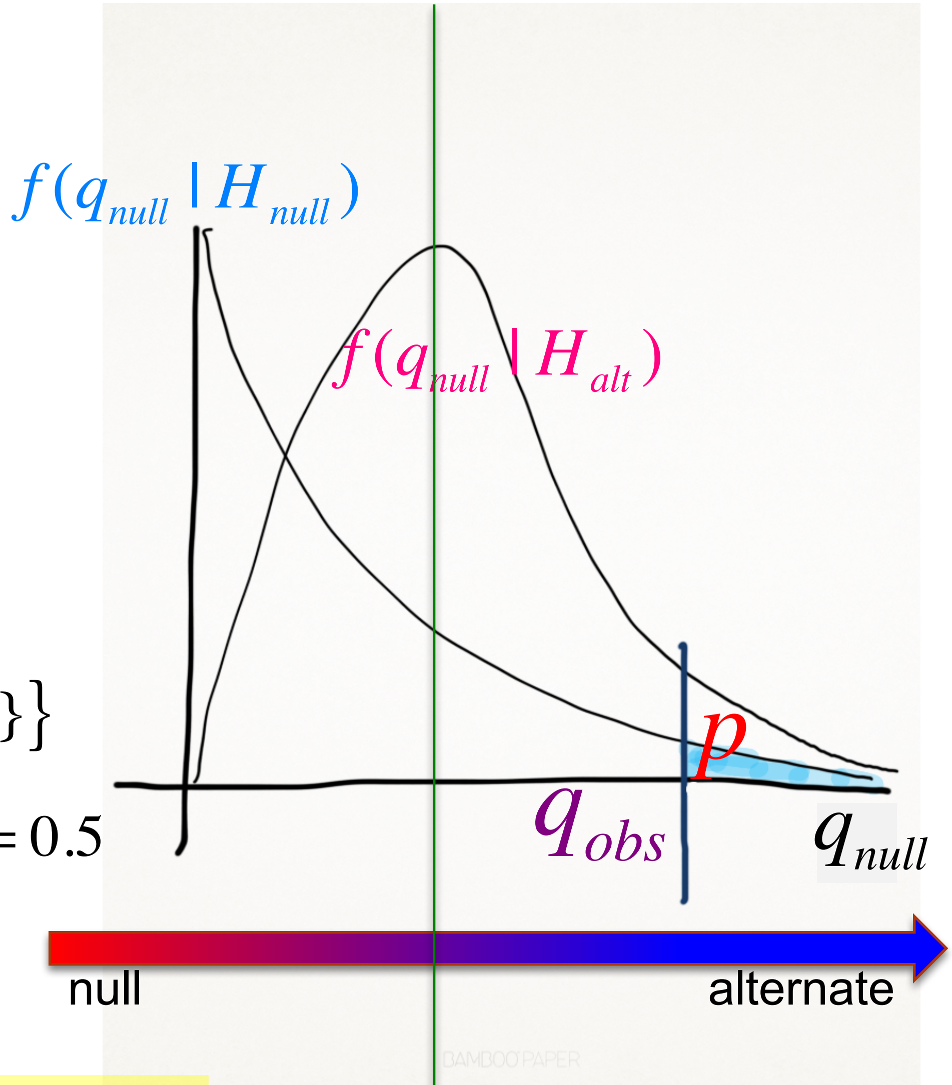
$f(q_{null} \mid H_{null})$

$q_{obs} \equiv q_{null,obs}$

$$p = \int_{q_{obs}}^{\infty} f(q_{null} \mid H_{null}) dq_{null}$$

$f(q_{null} \mid H_{alt})$

$$q_A \equiv q_{null,A} = \{q \mid med\{f(q_{null} \mid H_{alt})\}\} \qquad \int_{q_{null,A}}^{\infty} f(q_{null} \mid H_{null}) dq_{null} = 0.5$$



$$Z_{expected} = \sqrt{q_{null,A}}$$

$q_A \equiv q_{null,A}$

Resolving $f(q_{null} | H_{alt})$

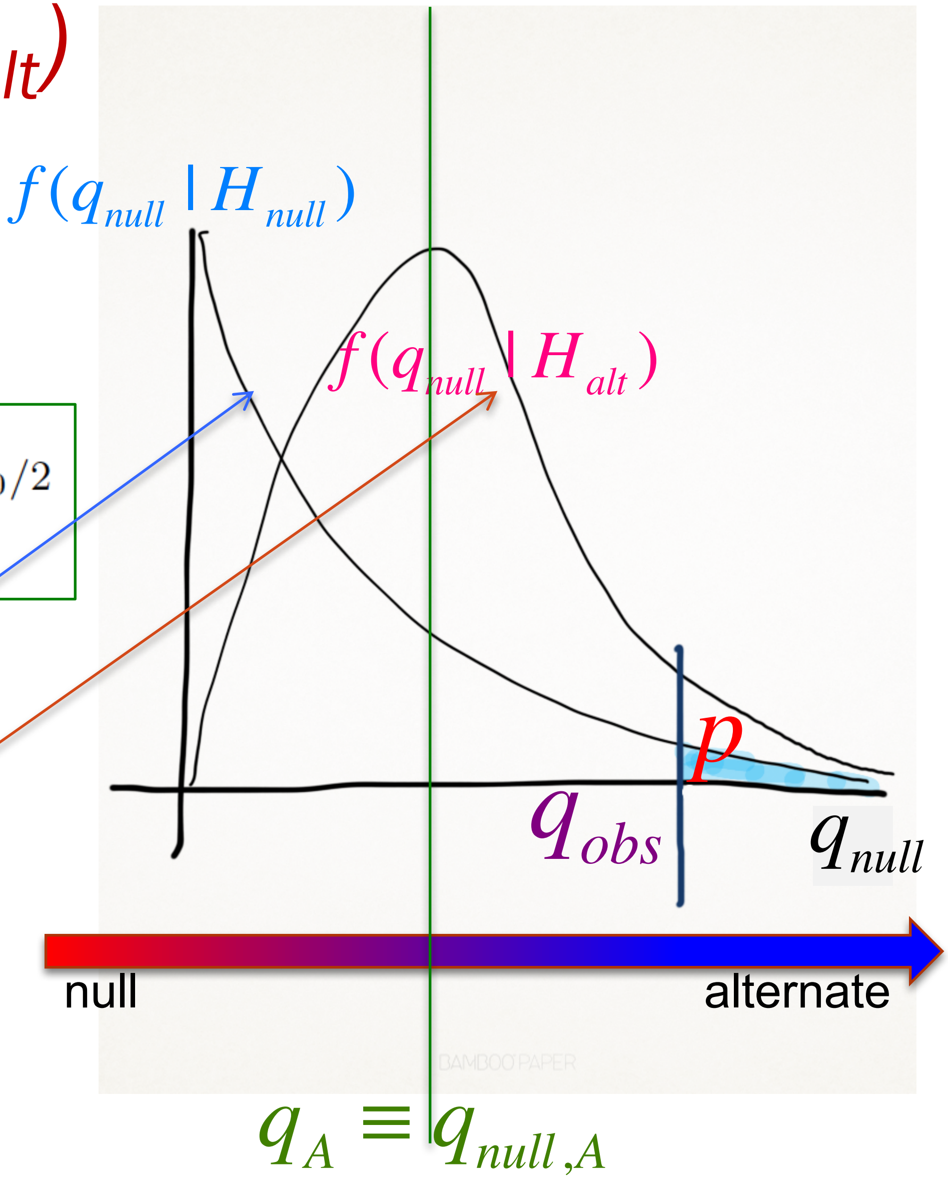
$$n = \mu s + b(\theta)$$

$$q_0 = \begin{cases} -2\ln\lambda(0) & \hat{\mu} \geq 0 \\ 0 & \hat{\mu} < 0 \end{cases}$$

$$f(q_0|0) = \frac{1}{2}\delta(q_0) + \frac{1}{2}\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{q_0}}e^{-q_0/2}$$

$$f(q_0 | 0) \sim \frac{1}{2}\chi^2$$

$$f(q_0 | \mu') \sim ?$$



Wald Theorem

- Consider a test of the strength parameter μ, which here can either be zero (for discovery) or nonzero (for an upper limit), and suppose the data are distributed according to a strength parameter μ′

$$\boxed{\langle \hat{\mu} \rangle = \mu'}$$

$$\lambda(\mu) = -2ln \frac{L(\mu s + \hat{\hat{b}}_{\mu})}{L(\hat{\mu} s + \hat{b})} = f(Data) = f(\hat{\mu})$$

- The desired distribution $f(q_{\mu} \mid \mu')$ can be found using a result due to Wald [1946], who showed that for the case of a single parameter of interest,

$$\boxed{-2\ln\lambda(\mu) = \frac{(\mu - \hat{\mu})^2}{\sigma^2} + \mathcal{O}(1/\sqrt{N})}$$

Wald Theorem

- Following the Wald Theorem we find that the 2-sided distributes like a non-central chi squared

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

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

$$f(t_\mu;\Lambda) = \frac{1}{2\sqrt{t_\mu}}\frac{1}{\sqrt{2\pi}}\left[\exp\left(-\frac{1}{2}\left(\sqrt{t_\mu}+\sqrt{\Lambda}\right)^2\right) + \exp\left(-\frac{1}{2}\left(\sqrt{t_\mu}-\sqrt{\Lambda}\right)^2\right)\right]$$

2 sided CI

$$\Lambda = \frac{(\mu-\mu')^2}{\sigma^2}$$

μ is the tested hypothesis while $\langle\hat{\mu}\rangle = \mu'$

under H_μ, if $\mu' = \mu$

we get Wilks theorem

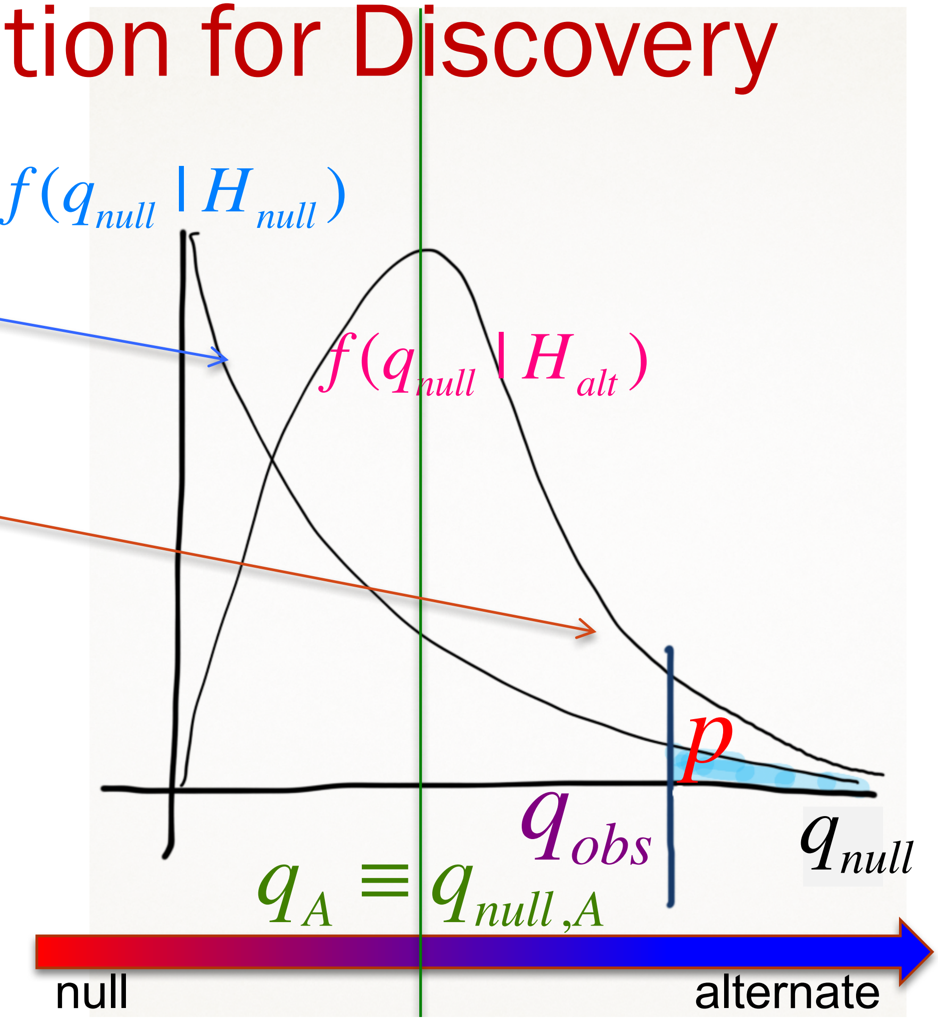
$$f(t_\mu|\mu) = \frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{t_\mu}}e^{-t_\mu/2}$$

The rediscovery of the Wald theorem helped to find the asymptotic distributions of all PL test Statistics, including the Neyman Pearson one, calculate the CLs modified p-values the expected sensitivity and save months if not years of computing

Asymptotic Distribution for Discovery

$$\boxed{f(q_0 \mid 0) \sim \frac{1}{2}\chi^2}$$

$$\boxed{f(q_0 \mid \mu') \sim ?}$$



1 sided CI

$$f(q_0|\mu') = \left(1 - \Phi\left(\frac{\mu'}{\sigma}\right)\right)\delta(q_0) + \frac{1}{2}\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{q_0}}\exp\left[-\frac{1}{2}\left(\sqrt{q_0} - \frac{\mu'}{\sigma}\right)^2\right]$$

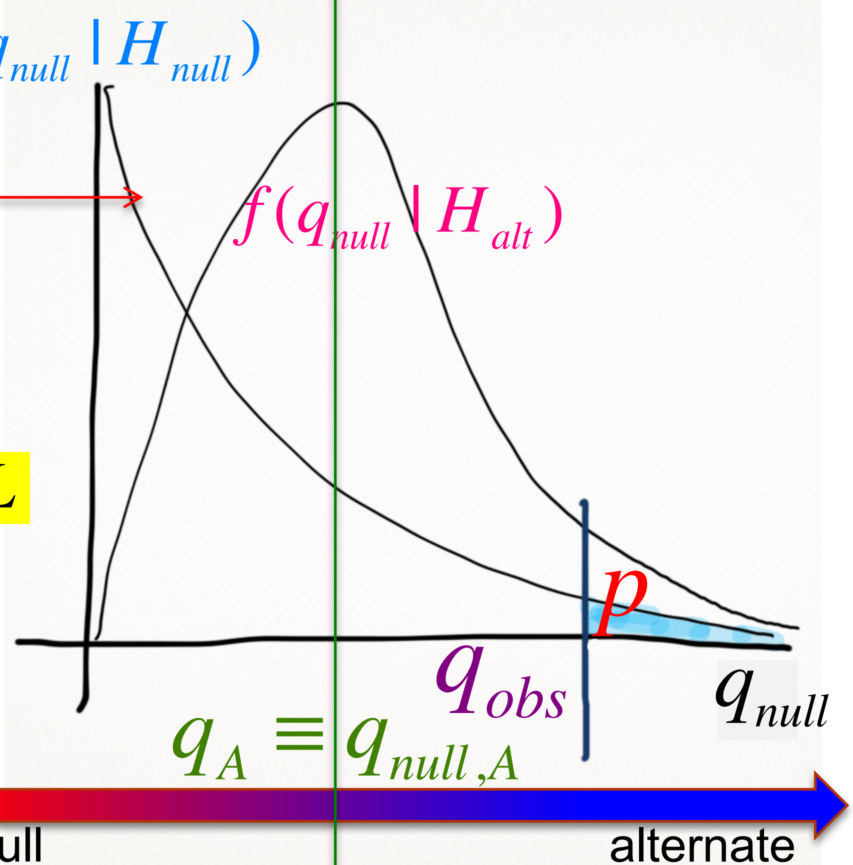
$$\Phi(Z) = 1 - \int_Z^\infty \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$$

Asymptotic Distribution for Exclusion

$$f(q_\mu|\mu) = \frac{1}{2}\delta(q_\mu) + \frac{1}{2}\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{q_\mu}}e^{-q_\mu/2}$$

$$\mu_{\rm up} = \hat{\mu} + \sigma\Phi^{-1}(1-\alpha)$$

$\alpha = 1 - CL$



1 sided CI

$$f(q_\mu|\mu') = \Phi\left(\frac{\mu'-\mu}{\sigma}\right)\delta(q_\mu) + \frac{1}{2}\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{q_\mu}}\exp\left[-\frac{1}{2}\left(\sqrt{q_\mu} - \frac{\mu-\mu'}{\sigma}\right)^2\right]$$

$$\Phi(Z) = 1 - \int_Z^\infty \frac{1}{\sqrt{2\pi}}e^{-x^2/2}dx$$

Asymptotic Distribution for FC

Depends on the observation
one might get 1-sided or 2-sided CI

3.4 Distribution of $\tilde{t}_\mu$

Assuming the Wald approximation, the statistic $\tilde{t}_\mu$ as defined by Eq. (11) can be written

$$\tilde{t}_\mu = \begin{cases} \frac{\mu^2}{\sigma^2} - \frac{2\mu\hat{\mu}}{\sigma^2} & \hat{\mu} < 0 \,, \\ \frac{(\mu-\hat{\mu})^2}{\sigma^2} & \hat{\mu} \geq 0 \,. \end{cases} \tag{40}$$

From this the pdf $f(\tilde{t}_\mu|\mu')$ is found to be

$$f(\tilde{t}_\mu|\mu') = \frac{1}{2}\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{\tilde{t}_\mu}}\exp\left[-\frac{1}{2}\left(\sqrt{\tilde{t}_\mu}+\frac{\mu-\mu'}{\sigma}\right)^2\right] \tag{41}$$

$$+ \begin{cases} \frac{1}{2}\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{\tilde{t}_\mu}}\exp\left[-\frac{1}{2}\left(\sqrt{\tilde{t}_\mu}-\frac{\mu-\mu'}{\sigma}\right)^2\right] & \tilde{t}_\mu \leq \mu^2/\sigma^2 \,, \\ \frac{1}{\sqrt{2\pi}(2\mu/\sigma)}\exp\left[-\frac{1}{2}\frac{\left(\tilde{t}_\mu-\frac{\mu^2-2\mu\mu'}{\sigma^2}\right)^2}{(2\mu/\sigma)^2}\right] & \tilde{t}_\mu > \mu^2/\sigma^2 \end{cases} \,. \tag{42}$$

The special case $\mu = \mu'$ is therefore

$$f(\tilde{t}_\mu|\mu') = \begin{cases} \frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{\tilde{t}_\mu}}e^{-\tilde{t}_\mu/2} & \tilde{t}_\mu \leq \mu^2/\sigma^2 \,, \\ \frac{1}{2}\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{\tilde{t}_\mu}}e^{-\tilde{t}_\mu/2} + \frac{1}{\sqrt{2\pi}(2\mu/\sigma)}\exp\left[-\frac{1}{2}\frac{(\tilde{t}_\mu+\mu^2/\sigma^2)^2}{(2\mu/\sigma)^2}\right] & \tilde{t}_\mu > \mu^2/\sigma^2 \,. \end{cases} \tag{43}$$

How to determine σ

- To estimate the uncertainty σ there are a few possibilities
  - Given the asymptotic formulae, fit the distribution of

$$f(q_{null} \,|\, H_{alt}) = f(q_{\mu} \,|\, \mu')$$

and extract σ

  - Implement the Wald formula to the Asimov data set and find

$$\sigma_A^2 = \frac{(\mu - \mu')^2}{q_{\mu,A}}$$

where μ is the tested (null) hypothesis and μ' is the alt hypothesis. For discovery, μ=0 while for exclusion μ'=0.

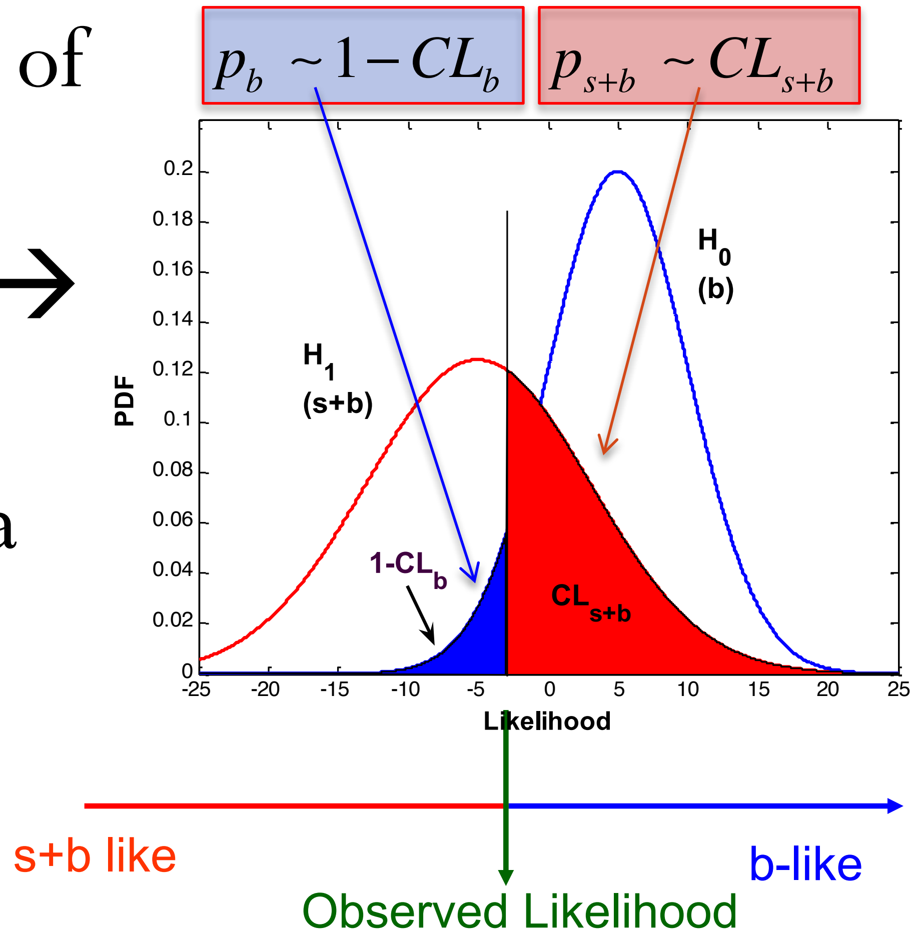
Exclusion and CLs

Case Study:

Exclusion of a Higgs with mass m_H

q_{NP} CL~p equivalence

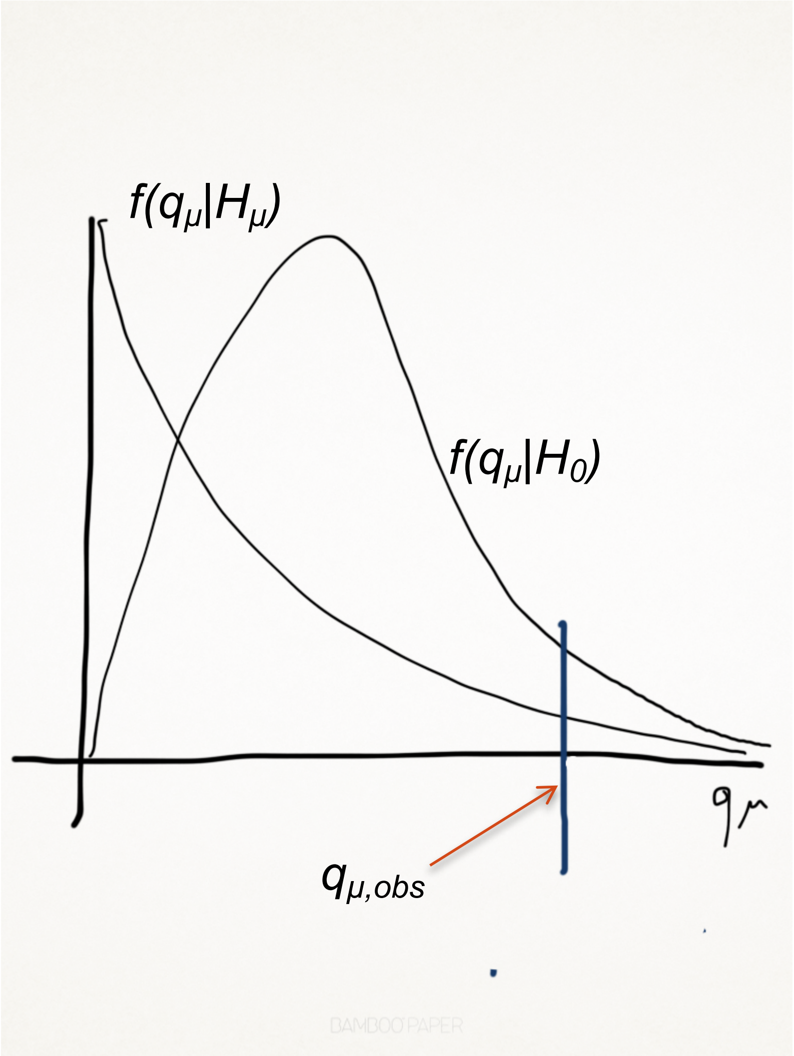
P_b is the incompatibility of the data with the background hypothesis → $CL_b \sim 1 - p_b$ is the compatibility of the data with the background hypothesis



March 2017

PL test statistics

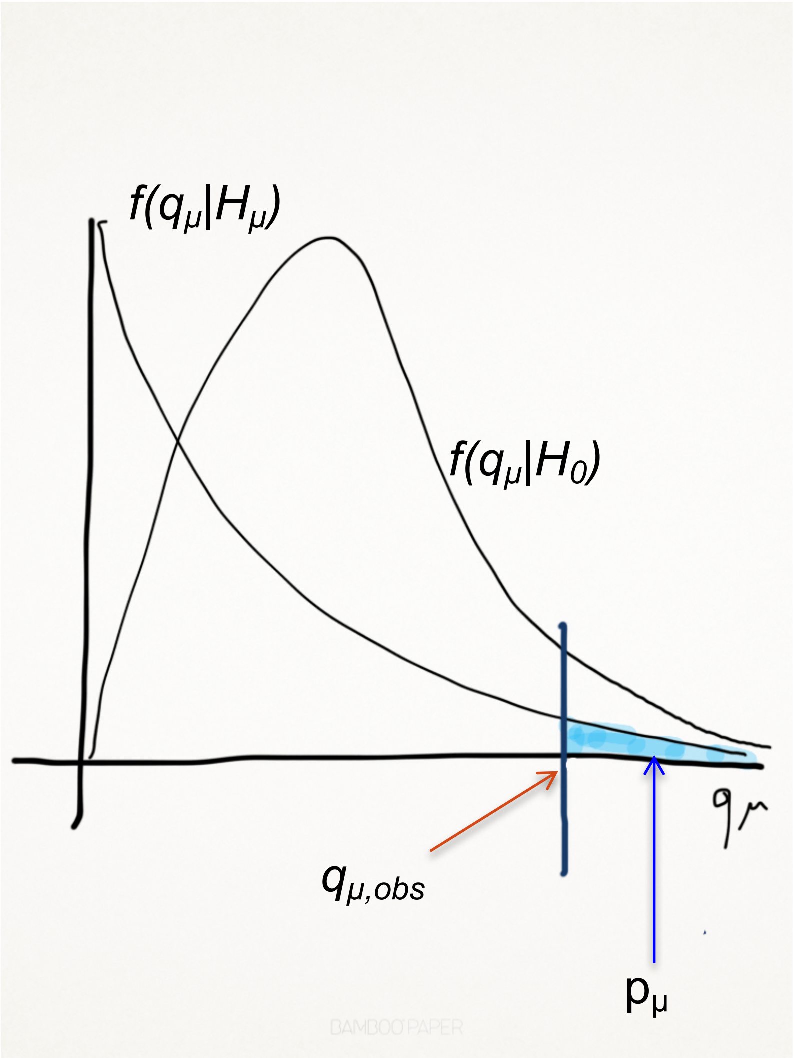
- We test hypothesis H_μ
- We calculate the PL (profile likelihood) ratio with the one observed data
- $q_{\mu,obs}$



- Find the p-value of the signal hypothesis H_μ

$$p_\mu = \int_{q_{\mu,obs}}^{\infty} f(q_\mu \mid H_\mu)\, dq_\mu$$

- In principle if $p_\mu<5\%$, H_μ hypothesis is excluded at the 95% CL
- Note that H_μ is for a given Higgs mass m_H



CLs

- Suppose $<n_b>=100$
- $s(m_{H1})=30$
- Suppose $n_{obs}=102$
- $s+b=130$
- $Prob(n_{obs}\le 102\,|\,130)<5\%$, m_{H1} is excluded at >95% CL

- Now suppose $s(m_{H2})=1$, can we exclude m_{H2}?
- Suppose $n_{obs}=80$, $prob(n_{obs}\le 80\,|\,101)<5\%$, it looks like we can exclude m_{H2}…
  but this is dangerous, because what we exclude is $(s(m_{H2})+b)$ and not s……
- With this logic we could also exclude b (expected b=100)
- To protect we calculate a modified p-value
- We cannot exclude m_{H2}

$$\frac{P(n\le n_o\,|\,s+b)}{P(n\le n_o\,|\,b)}=\frac{\mathrm{Prob}(\mathrm{nobs}\le 80\,|\,101)}{\mathrm{Prob}(\mathrm{nobs}\le 80\,|\,100)}\sim 1$$

CLs – the original derivation

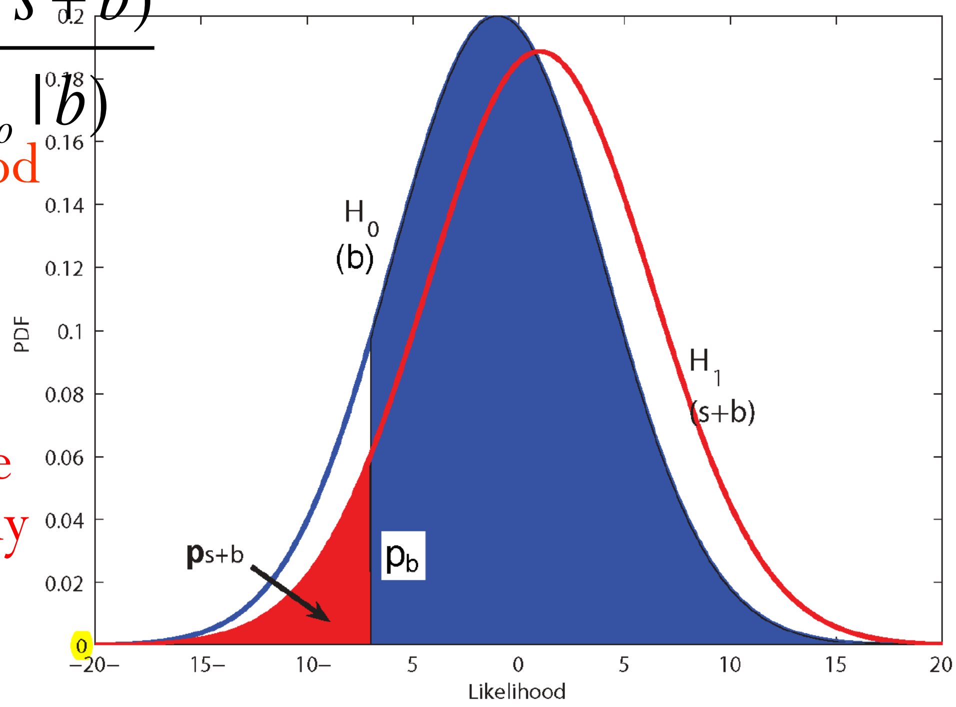
- Inspired by Zech derivation for counting experiments

$$P(n \le n_o | n_b \le n_o, s+b) = \frac{P(n \le n_o \mid s+b)}{P(n \le n_o \mid b)}$$

- A. Read suggested the $\mathrm{CL_s}$ method with

$$\boxed{CL_s = \frac{CL_{s+b}}{CL_b} = \frac{p_{s+b}}{1-p_b}}$$

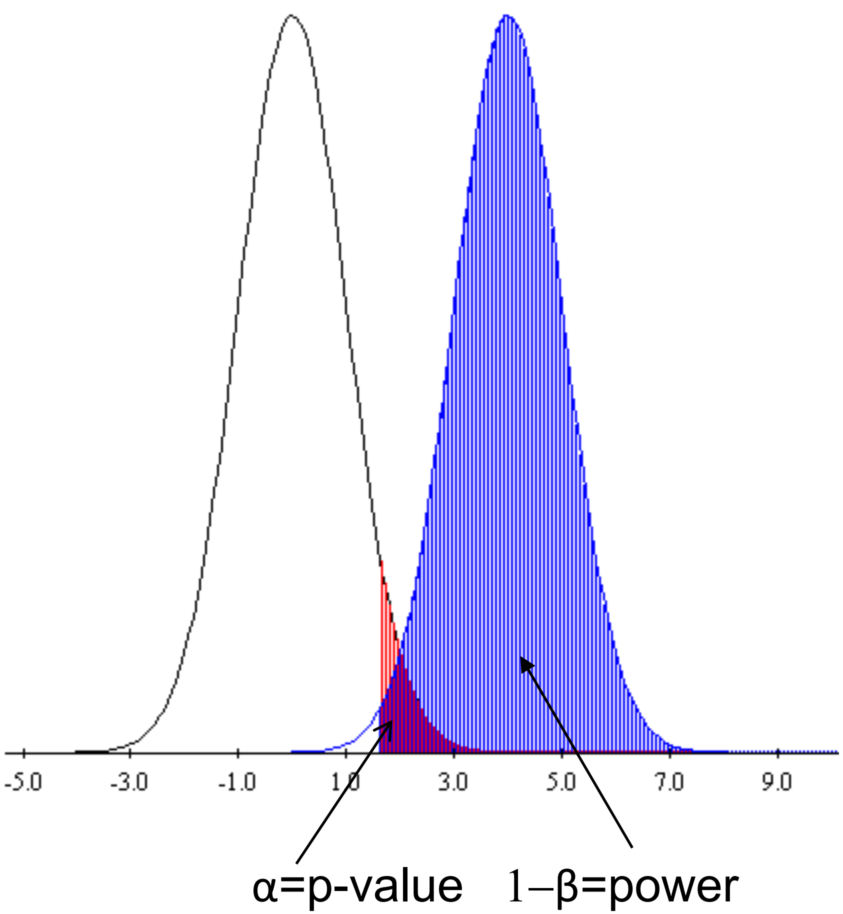
- This means that you will never be able to exclude a signal with a tiny cross section (to which you are not sensitive)



CLs Birenbaum 1962

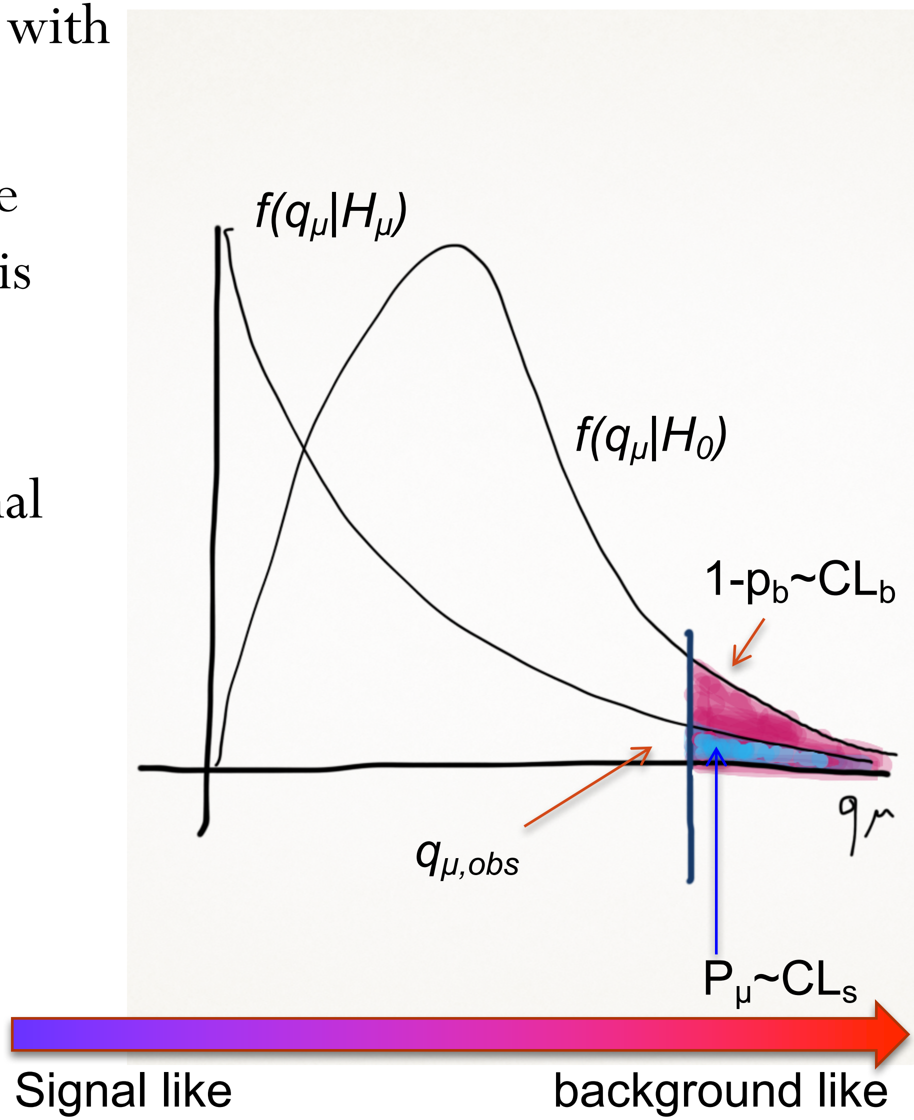
- Birnbaum suggested in 1962 that the the {p-value} / {power} should be used as a measure of the strength of statistical evidence provided by significance tests, rather than the {p-value} alone
- This translates into using a modified {p-value}

$$p'_s = \frac{p_{s+b}}{1 - p_b}$$



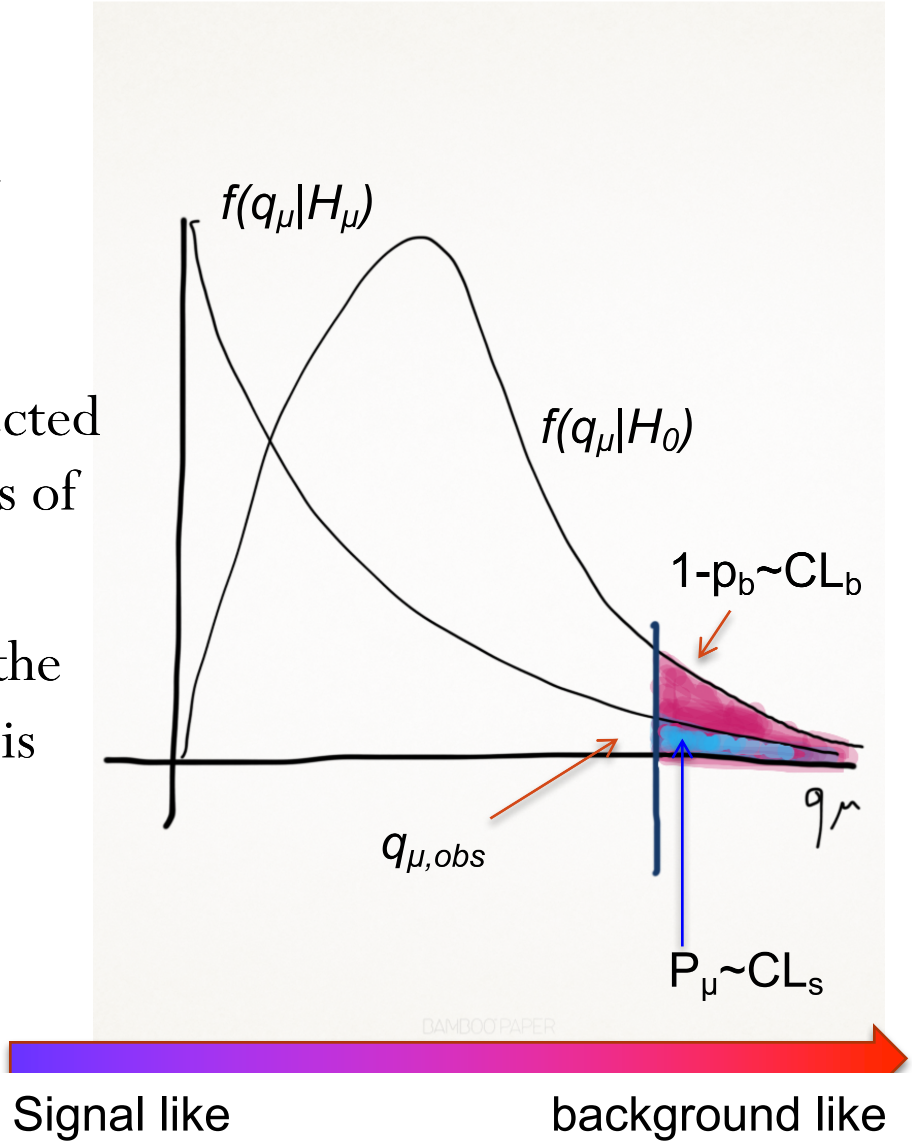
Second Verse- Same as the First (PL) CLb

- P_b is the incompatibility of the data with the background hypothesis →
- $CL_b \sim 1-p_b$ is the compatibility of the data with the background hypothesis and might be very small due to downward fluctuations of the background in the absence of a signal



CLs

- A complication arises when $\mu s+b \sim b$
- When the signal cross section is very small the $s(m_H)+b$ hypothesis can be rejected but at the same time the background hypothesis is almost rejected as well due to downward fluctuations of the background
- These downward fluctuations allow the exclusion of a signal the experiment is not sensitive to



The Modified CLs with the PL test statistic

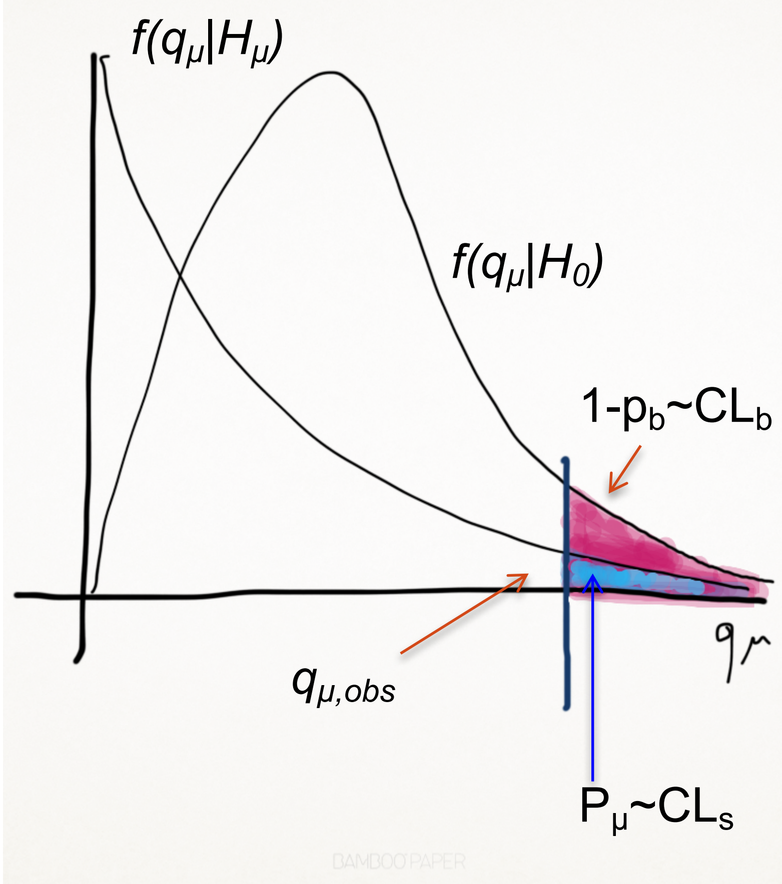
- The CLs method means that the signal hypothesis p-value

p_μ is modified to

$$p_\mu \to p'_\mu = \frac{p_\mu}{1-p_b}$$

$$p_\mu = \int_{\tilde{q}_{\mu,obs}}^{\infty} f(\tilde{q}_\mu|\mu)d\tilde{q}_\mu$$

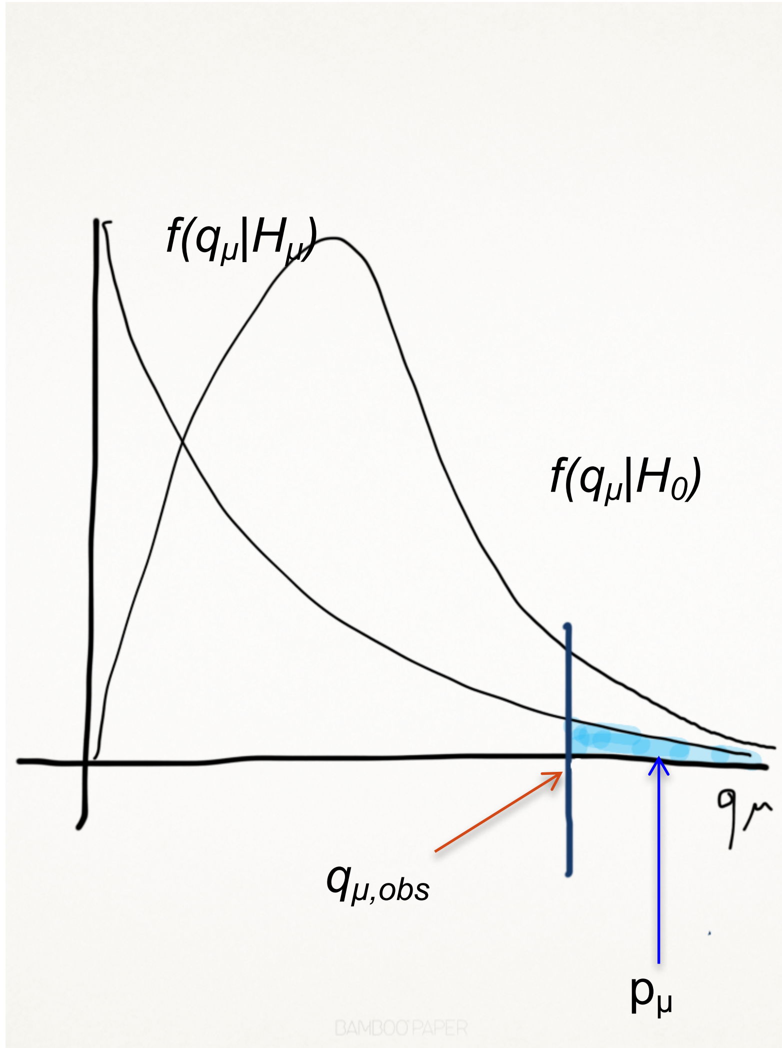
$$p_b = 1 - \int_{\tilde{q}_{\mu,obs}}^{\infty} f(\tilde{q}_\mu|0)d\tilde{q}_\mu$$



- Find the p-value of the signal hypothesis H_μ

$$p_\mu = \int_{q_{\mu,obs}}^{\infty} f(q_\mu \mid H_\mu) dq_\mu$$

- In principle if $p_\mu < 5\%$, H_μ hypothesis is excluded at the 95% CL
- Note that H_μ is for a given Higgs mass m_H



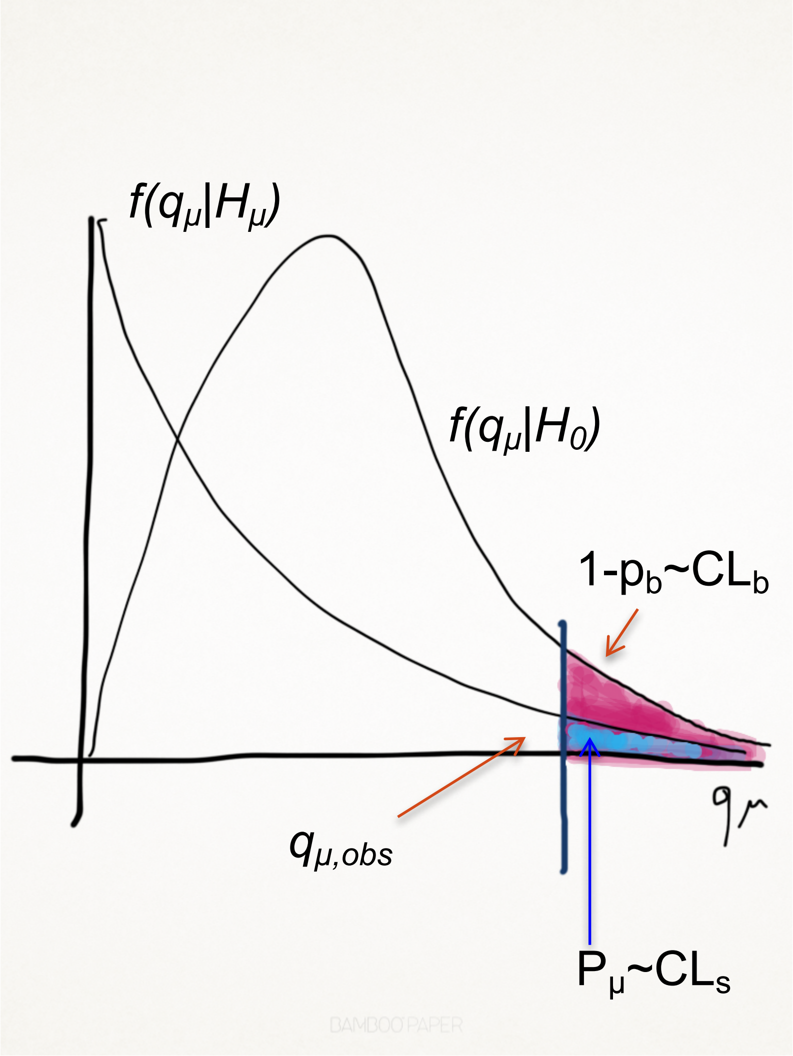
- Find the p-value of the signal hypothesis H_μ

- Find the modified p-value

$$p'_\mu(m_H) = \frac{p_\mu}{1-p_b}$$

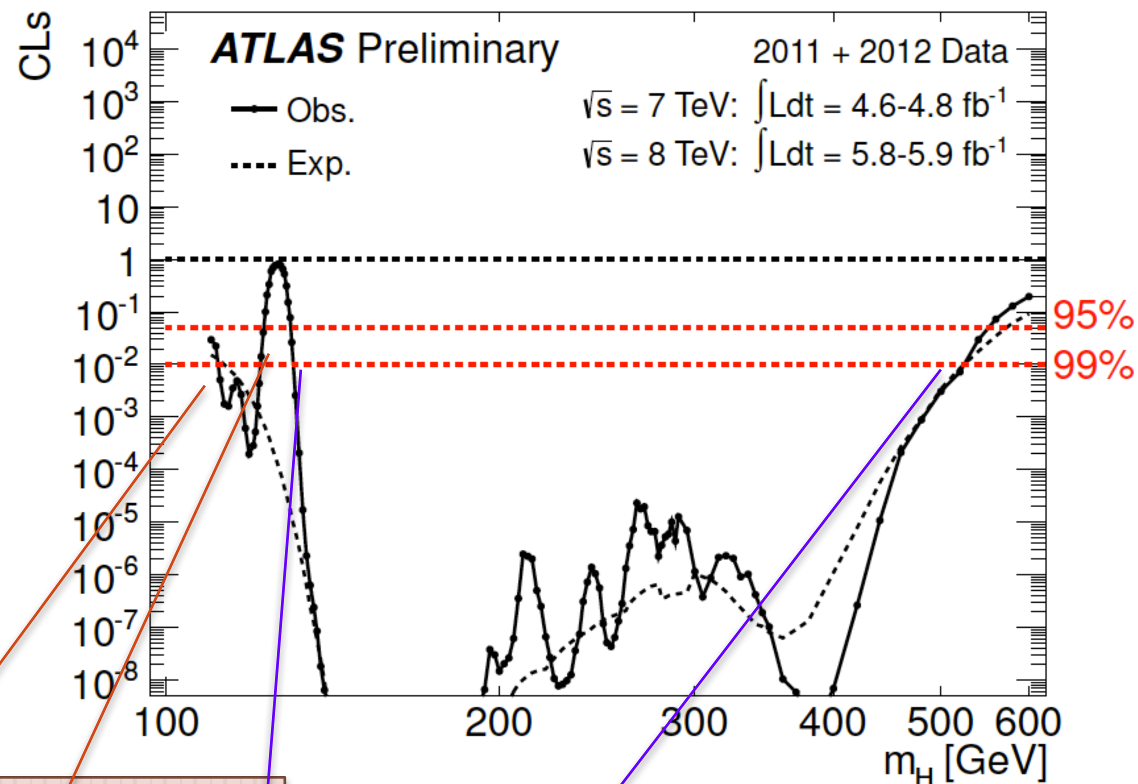
  - To tell if s is excluded, set $\mu=1$ and find

$$p'_1(m_H) = \frac{p_\mu}{1-p_b} \equiv CLs(m_H)$$



Understanding the CLs plot

- Here, for each Higgs mass m_H, one finds the observed p'_s value, i.e. $p'_{\mu},\mu=1$
- This modified p-value, p'_s, is by definition CLs

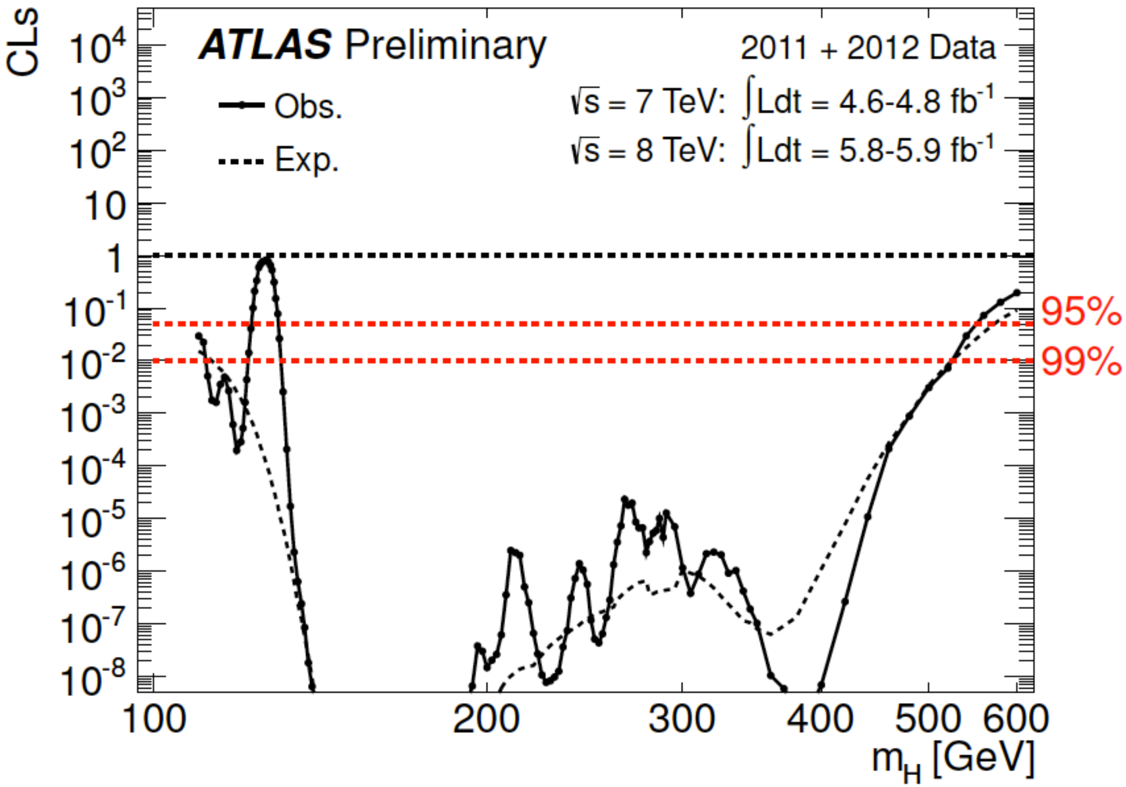


The smaller CLs, the deeper is the exclusion,
Exclusion CL=1-CLs=1-p'_s

to the previous combined search [1]. Figure 2 shows the CL_s values for $\mu = 1$, where it can be seen that the regions between 111.7 GeV to 121.8 GeV and 130.7 GeV to 523 GeV are excluded at the 99% CL.

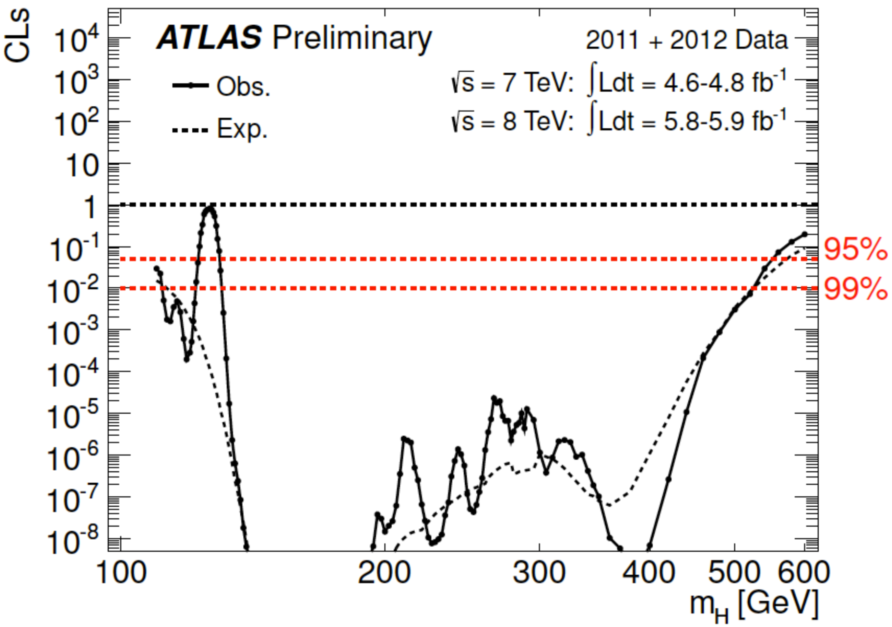
Understanding the CLs plot

- CLs is the compatibility of the data with the signal hypothesis
- The smaller the CLs, the less compatible the data with the prospective signal



- Find the p-value of the signal h... H_μ

$$p_\mu = \int_{q_{\mu,obs}}^{\infty} f(q_\mu \mid H_\mu)\, dq_\mu$$



- Find the modified p-value

$$p'_\mu(m_H) = \frac{p_\mu}{1 - p_b}$$

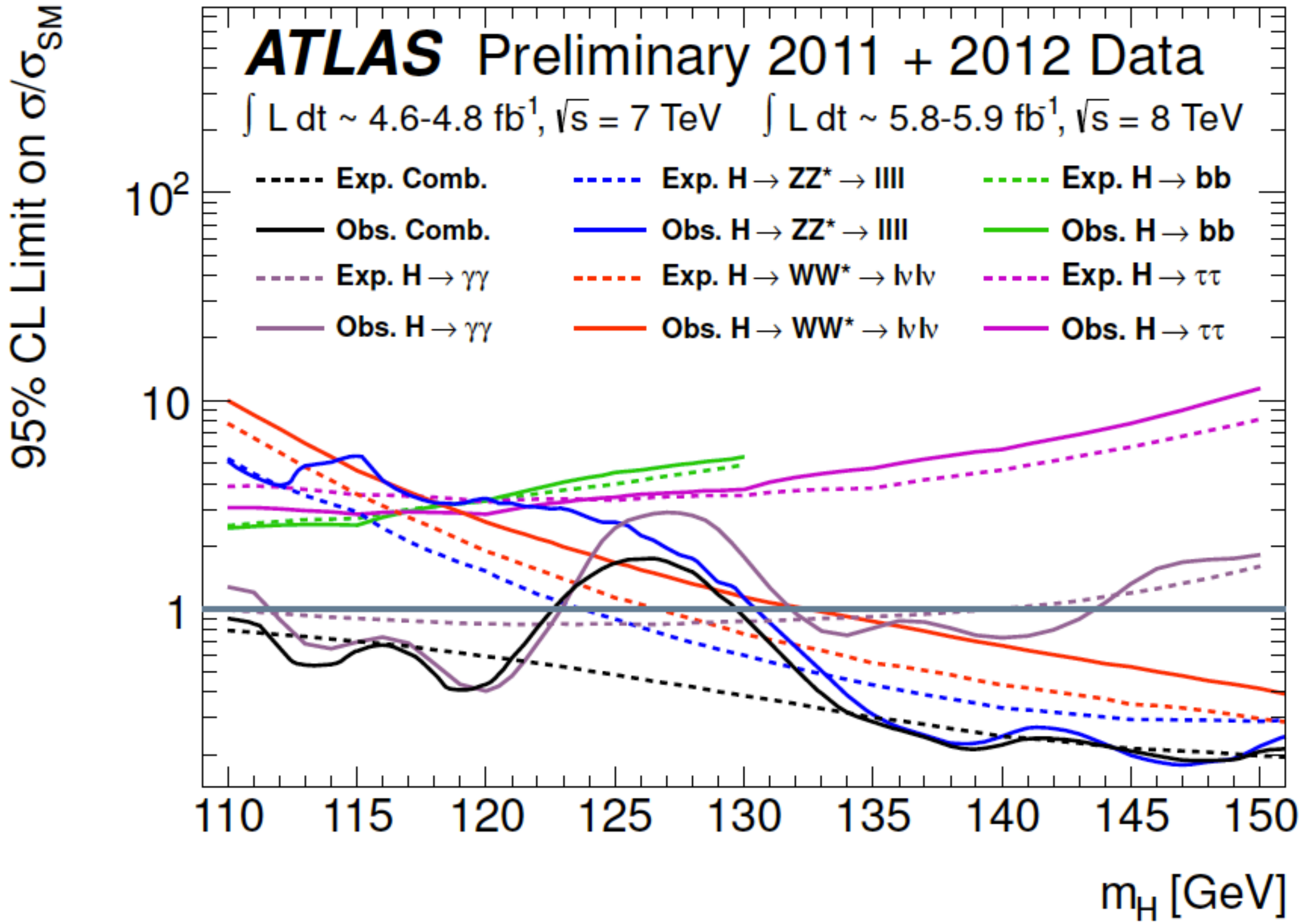
  - Option2: Iterate and find μ for which $p'_\mu(m_H)=5\%$ → μ=μup →
    If μup<1, m_H is excluded at the 95%

For a given data set, in the absence of a signal, the bigger the tested μ is the exclusion is deeper i.e. p'_μ is smaller

Exclusion a Higgs with a mass m_H

- First we fix the hypothesized mass to m_H
- We then test the H_μ $[\mu s(m_H)+b]$ hypothesis
- We find μ_{up}, for which $p'_{\mu up}=5\%$-> the $H_{\mu up}$ hypothesis is rejected at the 95% CL
- This means that the Confidence Interval of μ is $\mu \in [0,\mu_{up}]$
- If $\mu_{up}=\sigma(mH)/\sigma SM(mH)<1$, we claim that a SM Higgs with a mass m_H is excluded at the 95% CL
- A SM Higgs with a mass m_H, such that $\mu(m_H)<1$ is excluded at the 95% CL

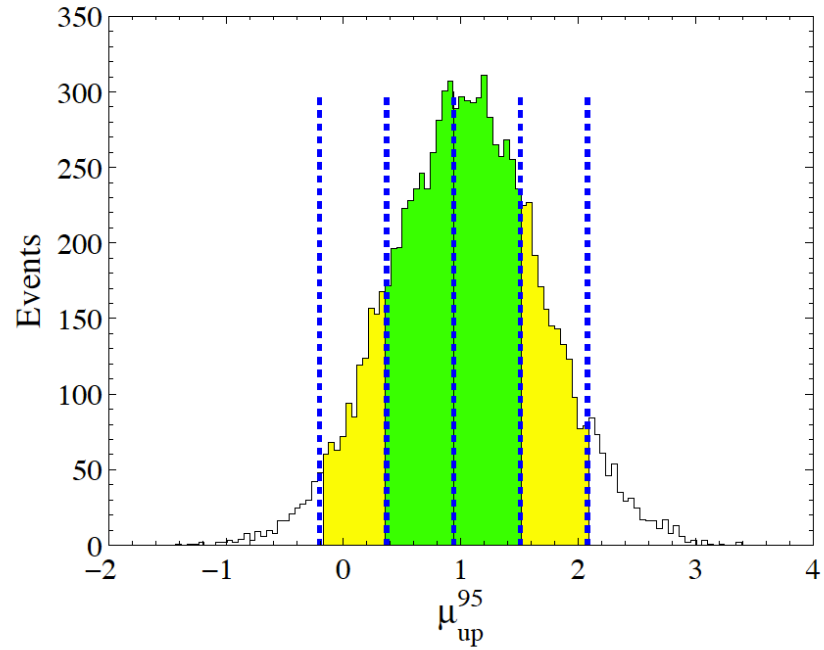
Upper Limit – $\mu_{up}(m_H)$



Sensitivity

- The sensitivity of an experiment to exclude a Higgs with a mass m_H is the median upper limit
- $$\mu_{up}^{med} = med\{\mu_{up} \mid H_0\}$$
- The 68% (green) and 95% (yellow) are the
  1 and 2 σ bands
- The median and the bands can be derived with the Asimov background only dataset n=b

Distribution of the upper limit with background only experiments



The Asimov data set is n=b
-> median upper limit

CCGV Useful Formulae – The Bands

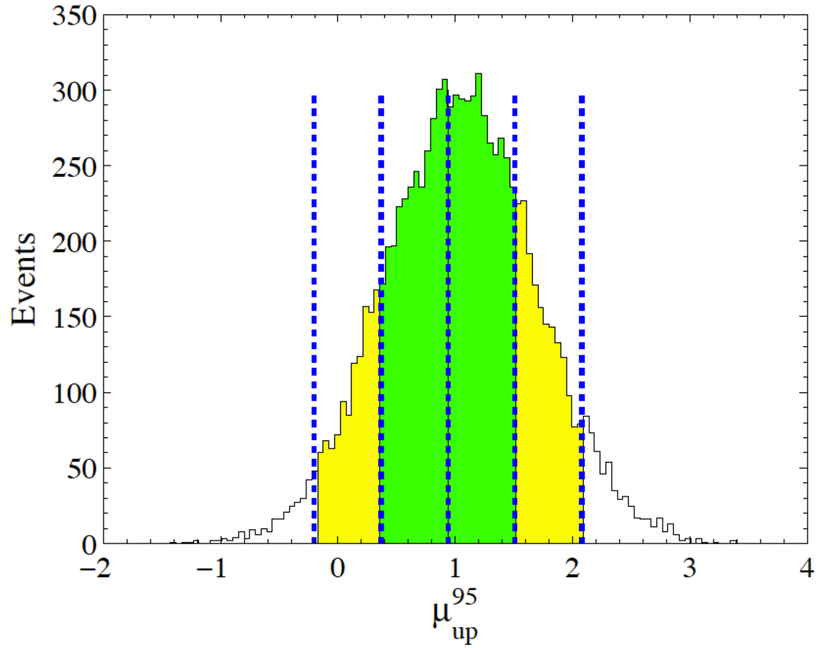
$$\mu_{up}^{med} = \hat{\mu} + \sigma_{\mu_{up}^{med}} \phi^{-1}(1-\alpha)$$

$$\alpha = 0.05 \rightarrow \phi^{-1}(1-\alpha) = \phi^{-1}(0.95) = 1.64$$

$$\mu_{up+N\sigma} = \hat{\mu} + \sigma_{up+N\sigma}\left(\phi^{-1}(1-\alpha) + N\right)$$

$$\sigma^2_{up+N\sigma} = \frac{\mu^2_{up+N\sigma}}{q_{up+N\sigma,A}}$$

Distribution of the upper limit with background only experiments



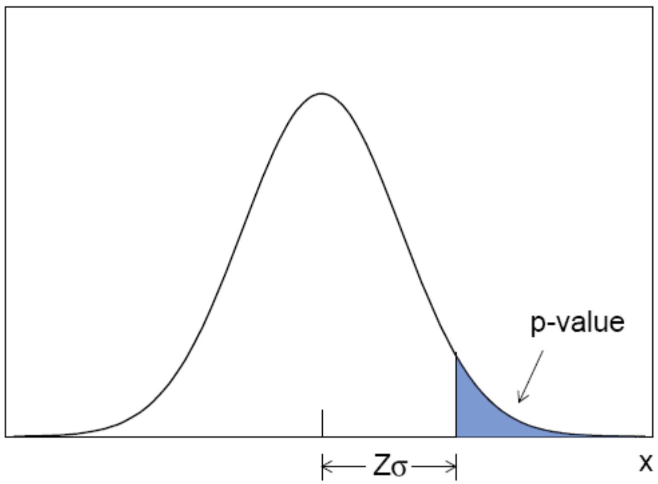
The Asimov data set is n=b
-> median upper limit

Useful Formulae

$$p'_{\mu_{95}} = \frac{1-\Phi(\sqrt{q_{\mu_{95}}})}{\Phi(\sqrt{q_{\mu_{95},A}} - \sqrt{q_{\mu_{95}}})} = 0.05$$

Φ is the cumulative distribution of the standard (zero mean, unit variance) Gaussian.

$q_{\mu_{95},A}$ Is evaluated with the Asimov data set (background only)



$$p = \int_Z^\infty \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, dx = 1 - \Phi(Z)$$

$$Z = \Phi^{-1}(1-p)$$

Understanding the Brazil Plot

The expected 95% CL exclusion region covers the m_H range from 110 GeV to 582 GeV. The observed 95% CL exclusion regions are from 110 GeV to 122.6 GeV and 129.7 GeV to 558 GeV. The addition of

- $\mu_{up}=\sigma(m_H)/\sigma_{SM}(m_H)<1 \rightarrow \sigma(m_H)<\sigma_{SM}(m_H) \rightarrow$ SM m_H excluded
- The line $\mu up=1$ corresponds to CLs=5% (p'_s=5%)
- The smaller $\mu up<1$ is, the exclusion of a SM Higgs is deeper$\rightarrow p's<5\%$, $p's=CLs \rightarrow CL=1-p's>95\%$