Statistical Combination:

definitions and conventions

Deliverables

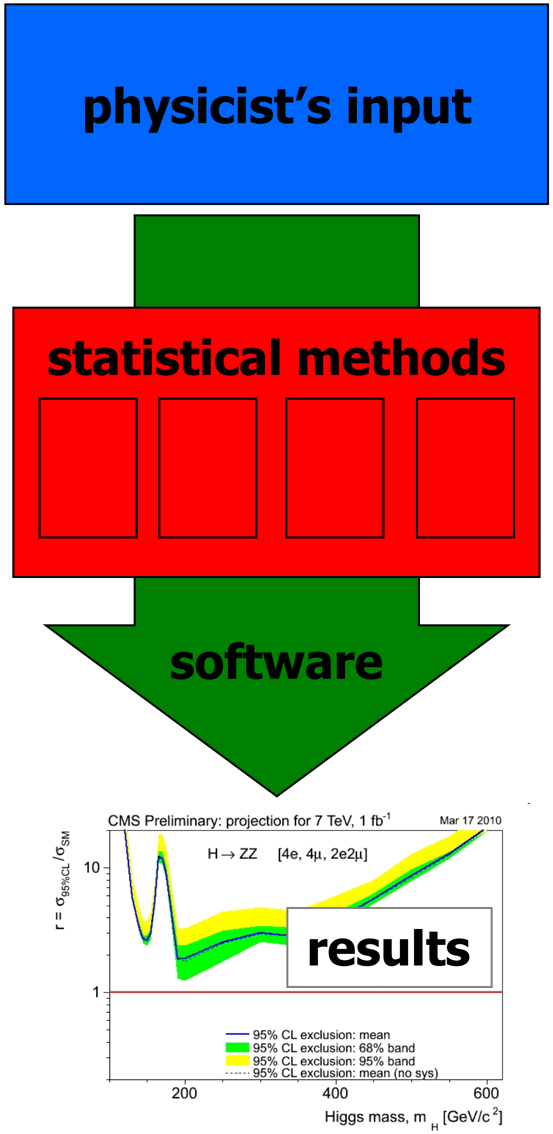
→ **statistical significance of an event excess**

→ **limits on the allowed signal strength**

→ **compatibility of the observation with expectations, both for bkgd-only and bkgd+signal hypotheses**

Do all of the above in a coherent and validated manner

Conventions



There are a number of conventions that we should agree on in order to be able to compare and combine

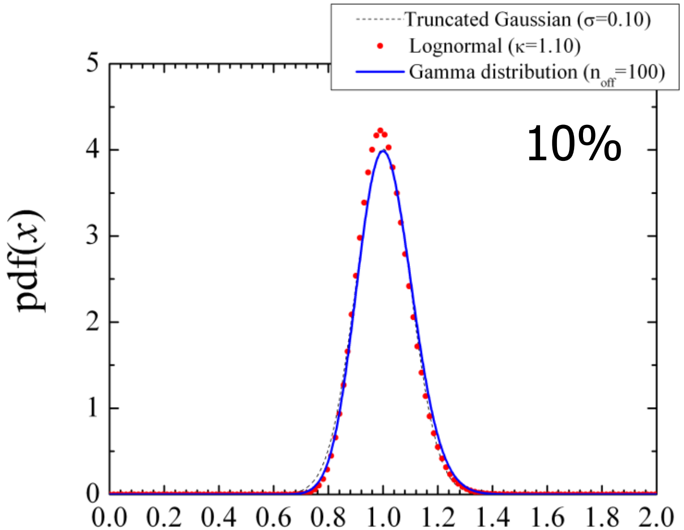
(1) pdf's for systematic errors at the input

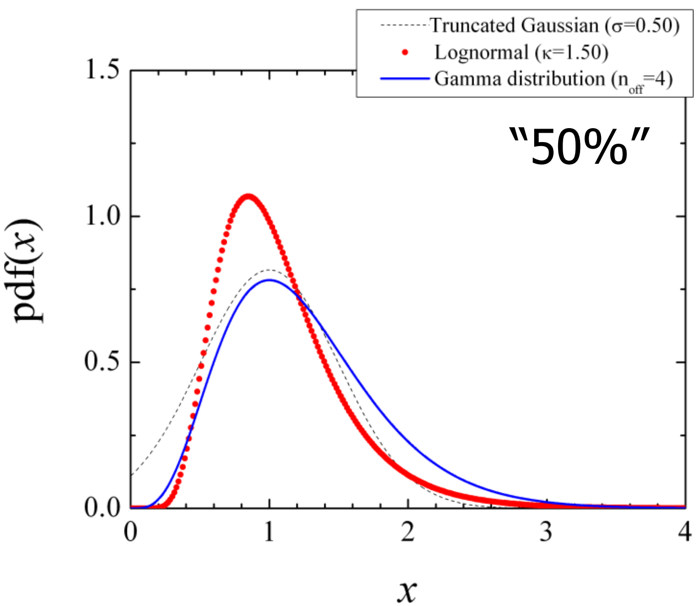
(2) given an observation, statistical methods for calculation of significance and exclusion limits

(3) quantifying expectations (mean or median, 68%/95%-bands)

As long as all conventions are well defined and followed, the rest is a matter of a technical execution

1: Systematic error pdf's





- **truncated Gaussian**
  not recommended for large errors >20%
  (unphysical and pathological in computations)

- **log-normal**
  recommended (identical to Gauss for small errors; physical and safe for large errors)

- **gamma-distribution**
  recommended when background is derived from a control sample: $n = \alpha \cdot N$ (identical to Gauss for small errors; physical and safe for large errors)

- **flat**
  OK, when justified

2(a) Exclusion limits

Bayesian

- prior on signal strength (flat, Jeffreys, reference, …)

Frequentist

- "Classical" (CL_{s+b})
- Modified Frequentist (CL_s)
- Power-constraint "Classical" Frequentist
- Each of the above frequentist methods has three sub-flavors:

$$Q = \frac{L(n \mid \mu s, b)}{L(n \mid b)} \qquad Q = \frac{L(n \mid \mu s, b; \theta)\big|_{\max(\theta)}}{L(n \mid b; \theta)\big|_{\max(\theta)}} \qquad Q = \frac{L(n \mid \mu s, b; \theta)\big|_{\max(\theta, \mu)}}{L(n \mid b; \theta)\big|_{\max(\theta)}}$$

Practical matters

- **pick one to be quoted in all abstracts**
- **pick one-two (more?) for comparisons to appear in the text**

2(a) Exclusion limits at Tevatron

Bayesian

- prior on signal strength (flat, Jeffreys, reference, …)

Frequentist

- "Classical" (CL_{s+b})
- Modified Frequentist (CL_s)
- Power-constraint "Classical" Frequentist
- Each of the above frequentist methods has three sub-flavors:

$$Q = \frac{L(n \mid \mu s, b)}{L(n \mid b)}$$

$$Q = \frac{L(n \mid \mu s, b; \theta)\big|_{\max(\theta)}}{L(n \mid b; \theta)\big|_{\max(\theta)}}$$

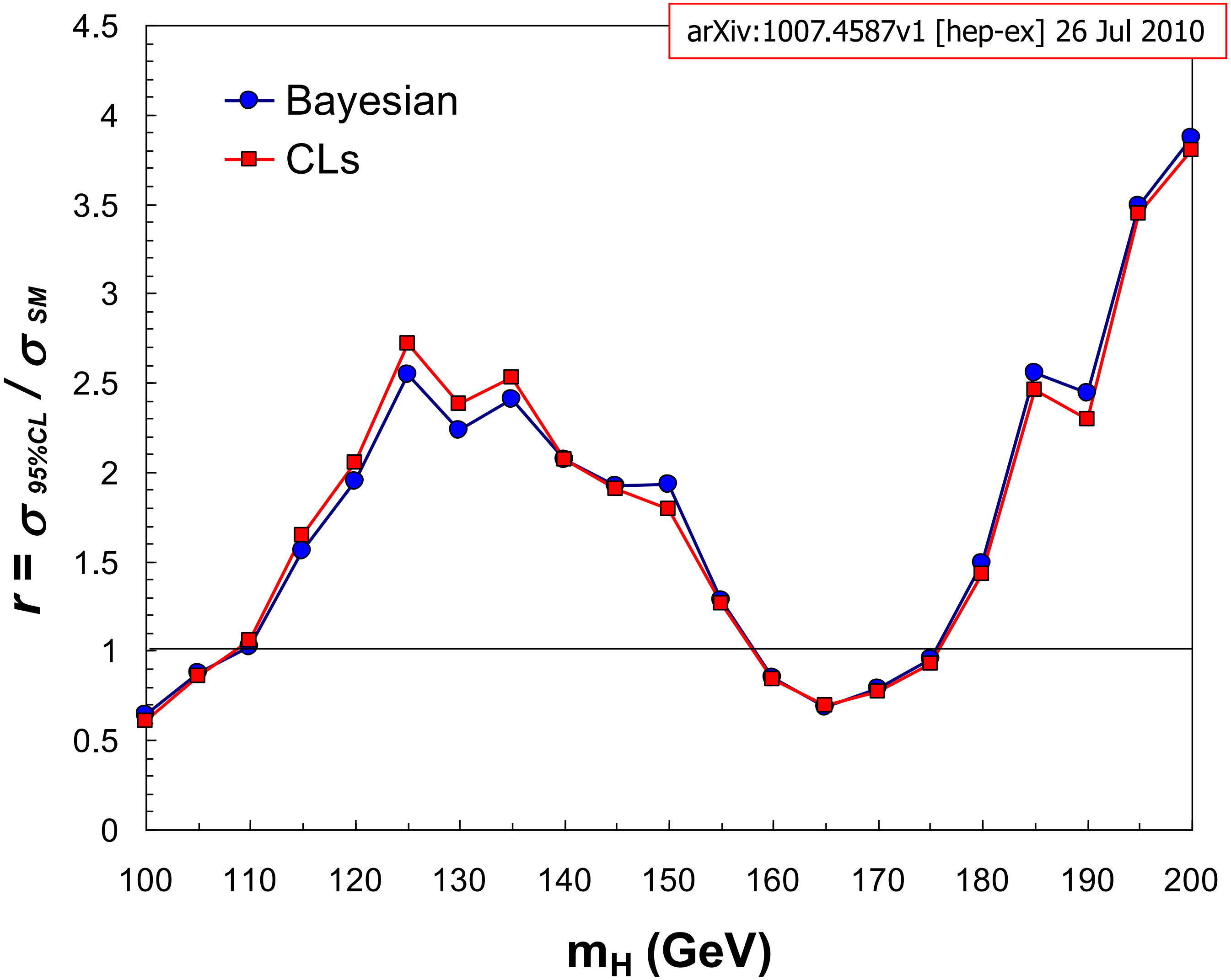
$$Q = \frac{L(n \mid \mu s, b; \theta)\big|_{\max(\theta, \mu)}}{L(n \mid b; \theta)\big|_{\max(\theta)}}$$

arXiv:1007.4587v1 [hep-ex] 26 Jul 2010

IV. COMBINING CHANNELS

To gain confidence that the final result does not depend on the details of the statistical formulation, we perform two types of combinations, using Bayesian and Modified Frequentist approaches, which yield limits on the Higgs boson production rate that agree within 10% at each value of m_H, and within 1% on average. Both methods rely on distributions in the final discriminants, and not just on their single integrated values. Systematic uncertainties enter on the predicted number of signal and background events as well as on the distribution of the discriminants in each analysis ("shape uncertainties"). Both methods use likelihood calculations based on Poisson probabilities.

2(a) Exclusion limits at Tevatron



2(b) Significance

Significance (Z) based on p-value

- Bayesian-Frequentist hybrid—intuitively natural
- evaluating p-value is CPU "expensive" for large values of Z (>5)

Profile Likelihood χ^2-approximation for large values of Z

- an approximation... but seems to work remarkably well for significance estimations in a wide range of initial settings

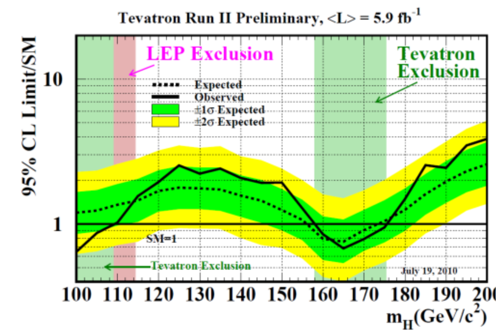
Quoting the scale of the look-elsewhere effect (a.k.a. trial factor)

- especially important for narrow peak searches in a wide range (makes a large impact at low values of Z<3)
- requires a priori definition of a search range

3: Expectation bands

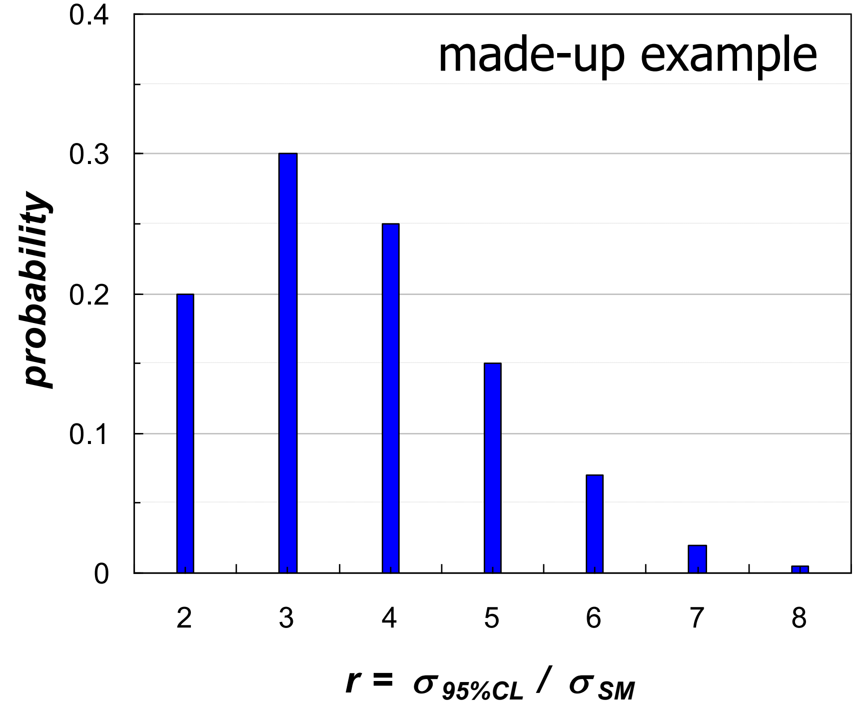
Expected limits are often represented by

- median, mean
- also, Azimov "typical" dataset for Bayesian limits
- 68%/95% (±1σ/±2σ) bands



For low statistics case,

- possible experimental outcomes are discrete
- median and 68%/95%-bands are subject to a convention



Summary

Summary of points subject to definitions and conventions is given

The actual definitions/conventions to be used in combining Higgs search results are yet to be chosen

As long as all conventions are well defined and followed, combination of search results is a matter of a technical execution