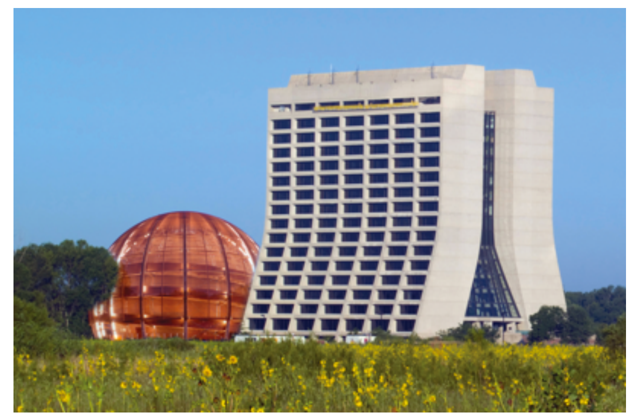
# Data Analysis and Statistical Methods in Experimental Particle Physics



Thomas R. Junk
*Fermilab*

Hadron Collider Physics Summer School 2012
August 6—17, 2012

# Lecture 4: Bayesian Inference, Binning, Smoothing

- Bayesian (re)definition of probability
- Handling Systematics
- Cross Section Measurements and Limits
- Fitting vs. Integrating

- Binning advice
- Density Estimation

# Reasons for Another Kind of Probability

- So far, we've been (mostly) using the notion that probability is the limit of a fraction of trials that pass a certain criterion to total trials.

- Systematic uncertainties involve many harder issues. Experimentalists spend much of their time evaluating and reducing the effects of systematic uncertainty.

- We also want more from our interpretations -- we want to be able to make decisions about what to do next.
  - Which HEP project to fund next?
  - Which theories to work on?
  - Which analysis topics within an experiment are likely to be fruitful?

These are all different kinds of bets that we are forced to make as scientists. They are fraught with uncertainty, subjectivity, and prejudice.

Non-scientists confront uncertainty and the need to make decisions too!

# Bayes' Theorem

Law of Joint Probability:

$$p(A \text{ and } B) = p(A|B)p(B) = p(B|A)p(A)$$

Events A and B interpreted to mean "data" and "hypothesis"

$$p(\{\nu\} \mid data) = \frac{L(data \mid \{\nu\})\pi(\nu)}{\int L(data \mid \{\nu'\})\pi(\{\nu'\})d\{\nu'\}}$$

$\{x\}$ = set of observations
$\{\nu\}$ = set of model parameters

A frequentist would say: Models have no "probability". One model's true, others are false. We just can't tell which ones (maybe the space of considered models does not contain a true one).
Better language: $p(\{\nu\} \mid data)$

describes our belief in the different models parameterized by $\{\nu\}$

# Bayes' Theorem

$p(\{\nu\} \mid data)$ is called the "posterior probability" of the model parameters

$\pi(\{\nu\})$ is called the "prior density" of the model parameters

The Bayesian approach tells us how our existing knowledge before we do the experiment is "updated" by having run the experiment.

This is a natural way to aggregate knowledge -- each experiment updates what we know from prior experiments (or subjective prejudice or some things which are obviously true, like physical region bounds).

Be sure not to aggregate the same information multiple times! (groupthink)

We make decisions and bets based on all of our knowledge and prejudices

"Every animal, even a frequentist statistician, is an informal Bayesian." See R. Cousins, "Why Isn't Every Physicist a Bayesian", Am. J. P., Volume 63, Issue 5, pp. 398-410

# How I remember Bayes's Theorem

$$p(\text{hypothesis}|\text{data}) = \frac{p(\text{data}|\text{hypothesis}) \times p(\text{hypothesis})}{p(\text{data})}$$

Posterior "PDF"
("Credibility")

"Likelihood Function"
("Bayesian Update")

"Prior belief
distribution"

Normalize this so that

$$\int p(\text{hypothesis}|\text{data}) d(\text{hypothesis}) = 1$$

for the observed data

# Bayesian Limits

Including uncertainties on nuisance parameters $\theta$

$$L'(data \mid r) = \int L(data \mid r, \theta)\pi(\theta)d\theta$$

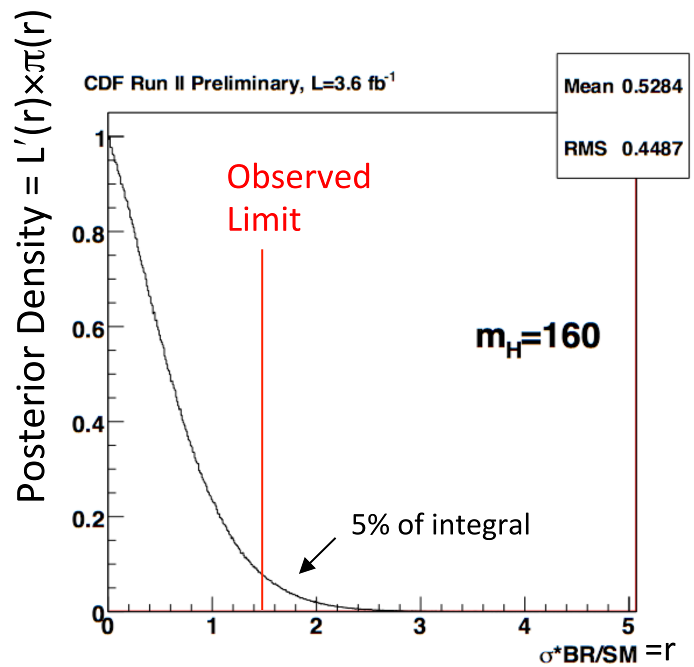
Typically $\pi(r)$ is constant
Other options possible.
**Sensitivity to priors a concern.**

where $\pi(\theta)$ encodes our prior belief in the values of the uncertain parameters. Usually Gaussian centered on the best estimate and with a width given by the systematic.
The integral is high-dimensional. Markov Chain MC integration is quite useful! Look up "Metropolis-Hastings Algorithm" on Wikipedia

Useful for a variety of results:

Limits:

$$0.95 = \frac{\int_0^{r_{\lim}} L'(data \mid r)\pi(r)dr}{\int_0^{\infty} L'(data \mid r)\pi(r)dr}$$



# Bayesian Cross Section Extraction

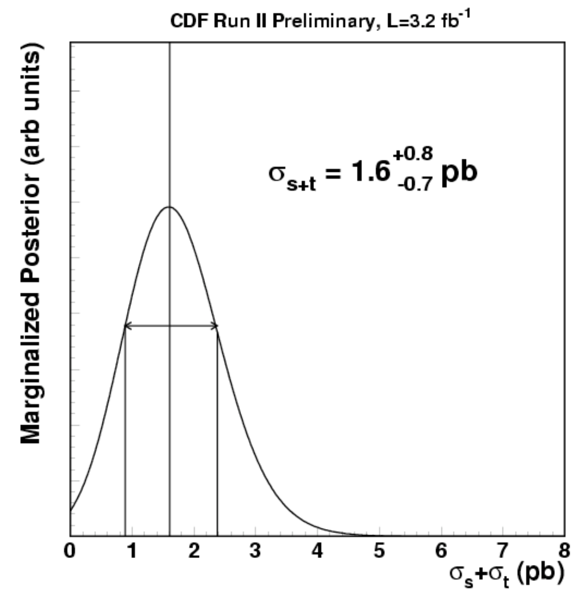
Same handling of nuisance parameters as for limits

$$L'(data\,|\,r) = \int L(data\,|\,r,\theta)\pi(\theta)d\theta$$

$$0.68 = \frac{\int_{r_{low}}^{r_{high}} L'(data\,|\,r)\pi(r)dr}{\int_{0}^{\infty} L'(data\,|\,r)\pi(r)dr}$$

The measured cross section and its uncertainty

$$r = r_{\max\,-(r_{\max}-r_{low})}^{\,+(r_{high}-r_{\max})}$$




Usually: shortest interval containing 68% of the posterior
(other choices possible). Use the word "credibility" in place of "confidence"

If the 68% CL interval does not contain zero, then the posterior at the top and bottom are equal in magnitude.
The interval can also break up into smaller pieces! (example: WW TGC@LEP2

# Extending Our Useful Tip About Limits

It takes almost exactly 3 expected signal events to exclude a model.

If you have zero events observed, zero expected background, and no systematic uncertainties, then the limit will be 3 signal events.

Call s=expected signal, b=expected background. r=s+b is the total prediction.

$$L(n=0,r) = \frac{r^0 e^{-r}}{0!} = e^{-r} = e^{-(s+b)}$$

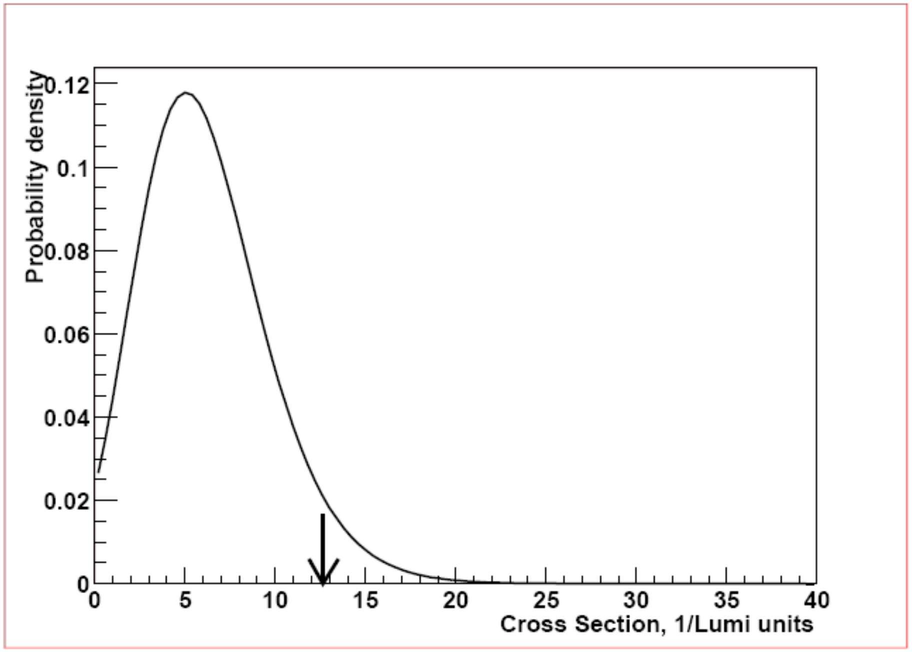
$$0.95 = \frac{\int_0^{r_{\lim}} L'(data \mid r)\pi(r)dr}{\int_0^{\infty} L'(data \mid r)\pi(r)dr} = \frac{-e^{-(s+b)}\Big|_0^{r_{\lim}}}{-e^{-(s+b)}\Big|_0^{\infty}} = e^{-r_{\lim}}$$

The background rate cancels! For 0 observed events, the signal limit does not depend on the predicted background (or its uncertainty). This is also true for $CL_s$ limits, but not PCL limits (which get stronger with more background)

If p=0.05, then r=-ln(0.05)=2.99573

# A Handy Limit Calculator

D0 (http://www-d0.fnal.gov/Run2Physics/limit_calc/limit_calc.html) has a web-based, menu-driven Bayesian limit calculator for a single counting experiment, with uncorrelated uncertainties on the acceptance, background, and luminosity. Assumes a uniform prior on the
signal strength. Computes 95% CL limits ("Credibility Level")




Data: 10
Background: 5 +- 1
Efficiency: 1.0 +- 0.1
Luminosity: 1.0 +- 0.0

The cross section 95% CL upper limit is 12.666

# Systematic Uncertainties

Encoded as priors on the nuisance parameters $\pi(\{\theta\})$.

Can be quite contentious -- injection of theory uncertainties and results from other experiments -- how much do we trust them?

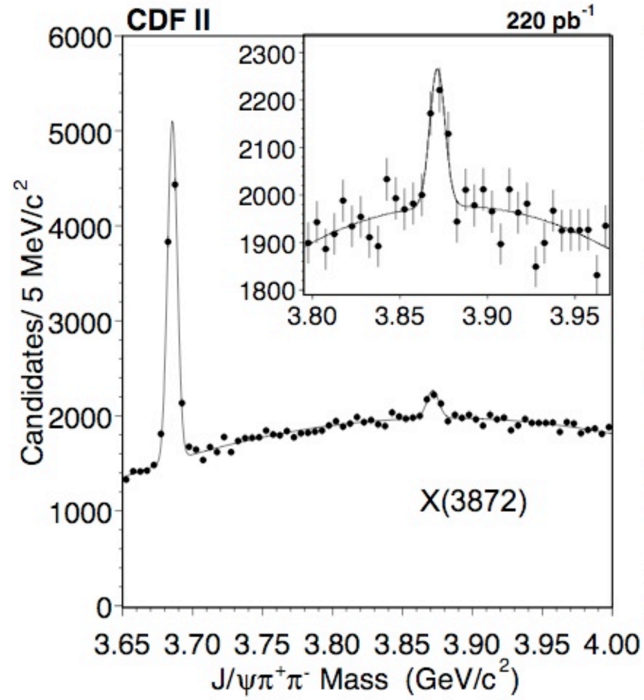
Do not inject the same information twice.

Some uncertainties have statistical interpretations -- can be included in L as additional data. Others are purely about belief. Theory errors often do not have statistical interpretations.

# Integrating over Systematic Uncertainties Helps Constrain their Values with Data

$$L'(data \mid r) = \int L(data \mid r,\theta)\pi(\theta)d\theta$$

Nuisance parameters: $\theta$
Parameter of Interest: $r$



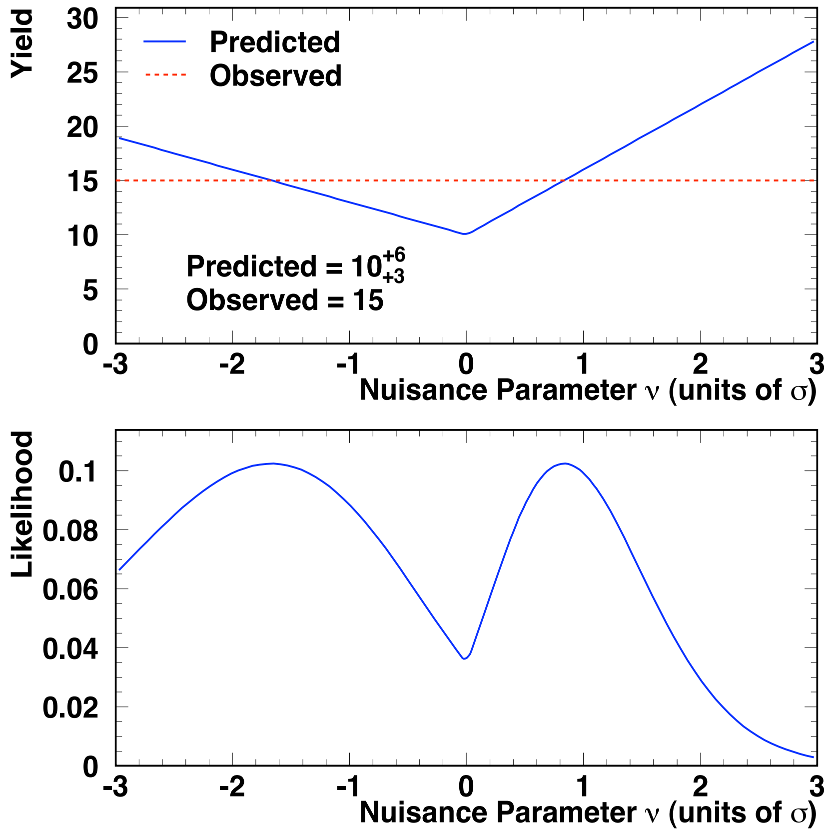

Example: suppose we have a background rate prediction that's 50% (fractionally) uncertain -- goes into $\pi(\theta)$. But only a narrow range of background rates contributes significantly to the integral. The kernel falls to zero rapidly outside of that range.

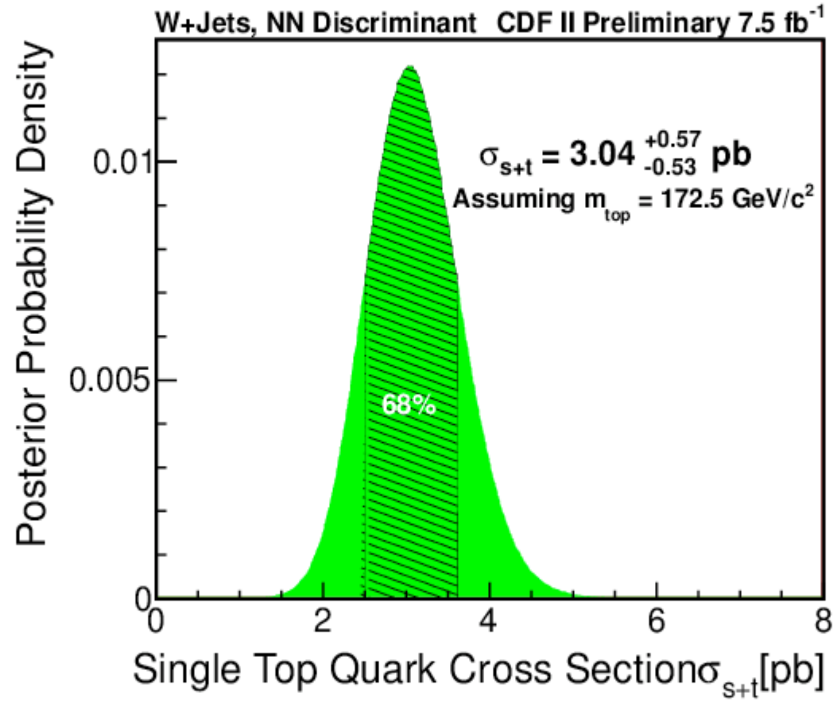
Can make a posterior probability distribution for the background too -- narrow belief distribution.

# Coping with Systematic Uncertainty

- "Profile:"
  - Maximize *L* over possible values of nuisance parameters include prior belief densities as part of the $\chi^2$ function (usually Gaussian constraints)
- "Marginalize:"
  - Integrate L over possible values of nuisance parameters (weighted by their prior belief functions -- Gaussian, gamma, others...)
  - Consistent Bayesian interpretation of uncertainty on nuisance parameters
- Aside: MC "statistical" uncertainties are systematic uncertainties

# Parameter Estimation – Marginalize or Profile?




If Pred = $10^{-6}_{-3}$, and obs=15, then the likelihood would have one maximum, but it would have a corner. MINUIT may quote inappropriate uncertainties as the second derivative isn't well defined.

The corner can be smoothed out – See
R. Barlow, http://arxiv.org/abs/physics/0406120,
http://arxiv.org/abs/physics/0401042
http://arxiv.org/abs/physics/0306138

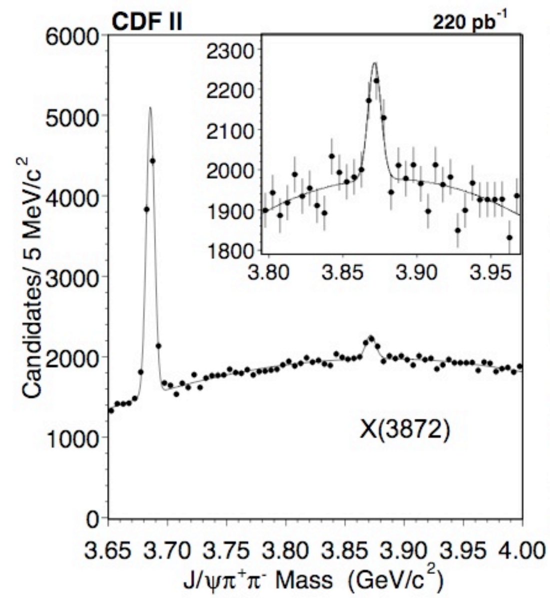
But I know of no way
to get rid of the double-peak
Nor should there be a way --
it can be a real effect. See the LEP2 TGC measurements

# Asymmetric Uncertainties and Priors

Measurements, and even theoretical calculations, frequently are assigned asymmetric uncertainties:

Value = $10^{+2}_{-1}$, or more extremely, $10^{+2}_{+2}$ (ouch). When the uncertainties have the same sign on both sides, it is worthwhile to check and see why this is the case.

Example – we seek a bump in a mass distribution by counting events in a small window around where the bump is sought.




The detector calibration has an energy uncertainty (magnetic field or chamber alignment for tracks, or much larger effect, calorimeter energy scales for jets).

Shift the calibration scale up – predicted peak shifts out of the window → downward shift in expected signal prediction.

Shift the calibration down – predicted peak shifts out of the other side of the window → downward shift in expected signal prediction

# Treatment of Asymmetric Uncertainties

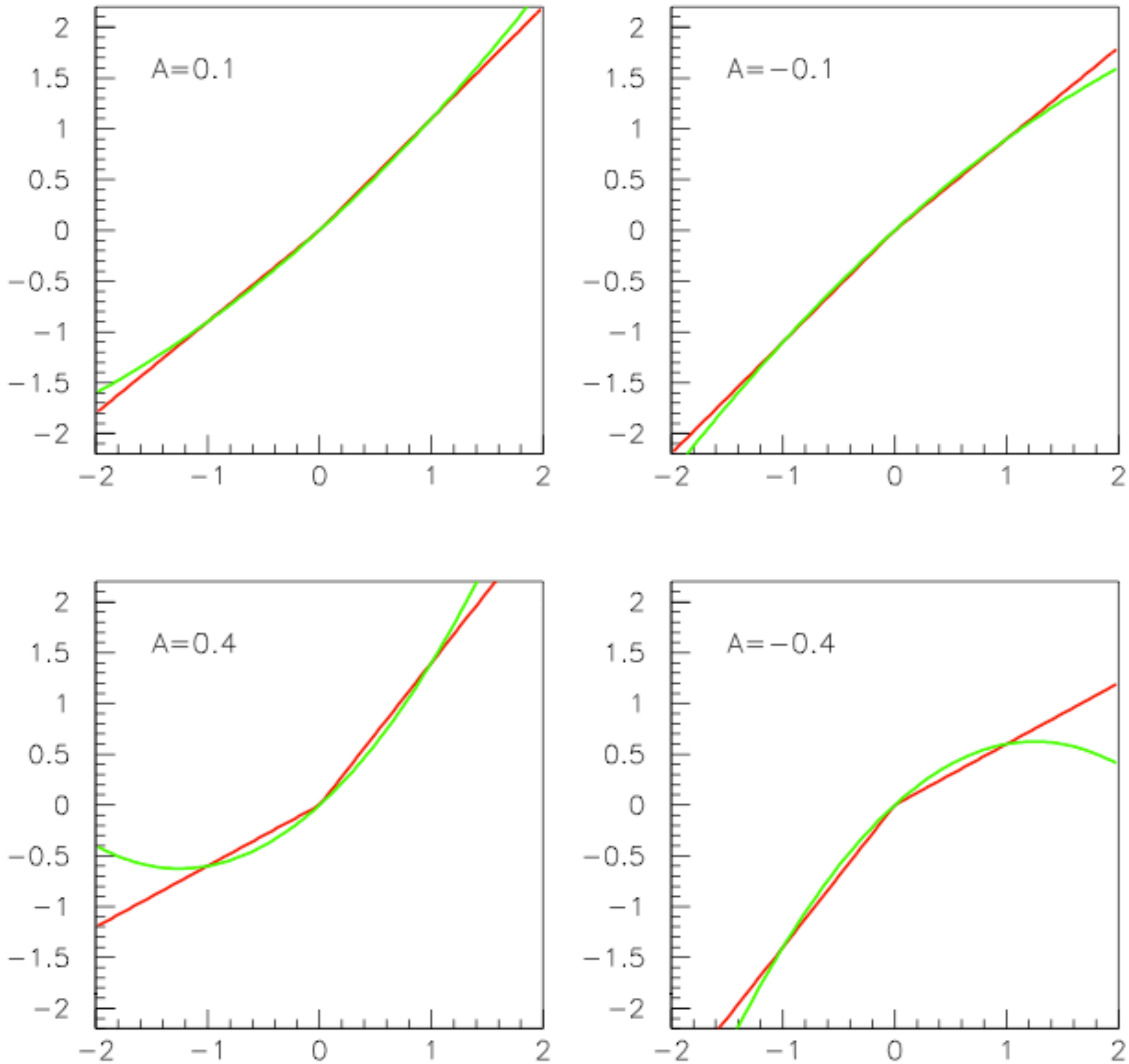
These cases are pretty clear – the underlying parameter, the energy scale, has a (Gaussian? Your choice) distribution, while it has a nonlinear, possibly non-monotonic *impact* on the model prediction.

The same parameter may have a linear, symmetrical impact on another model prediction, and we will have to treat them as correlated in statistical analysis tools.

Treatment is ambiguous when little is known why the uncertainties are asymmetric, or it is not clear how to extrapolate/interpolate them.

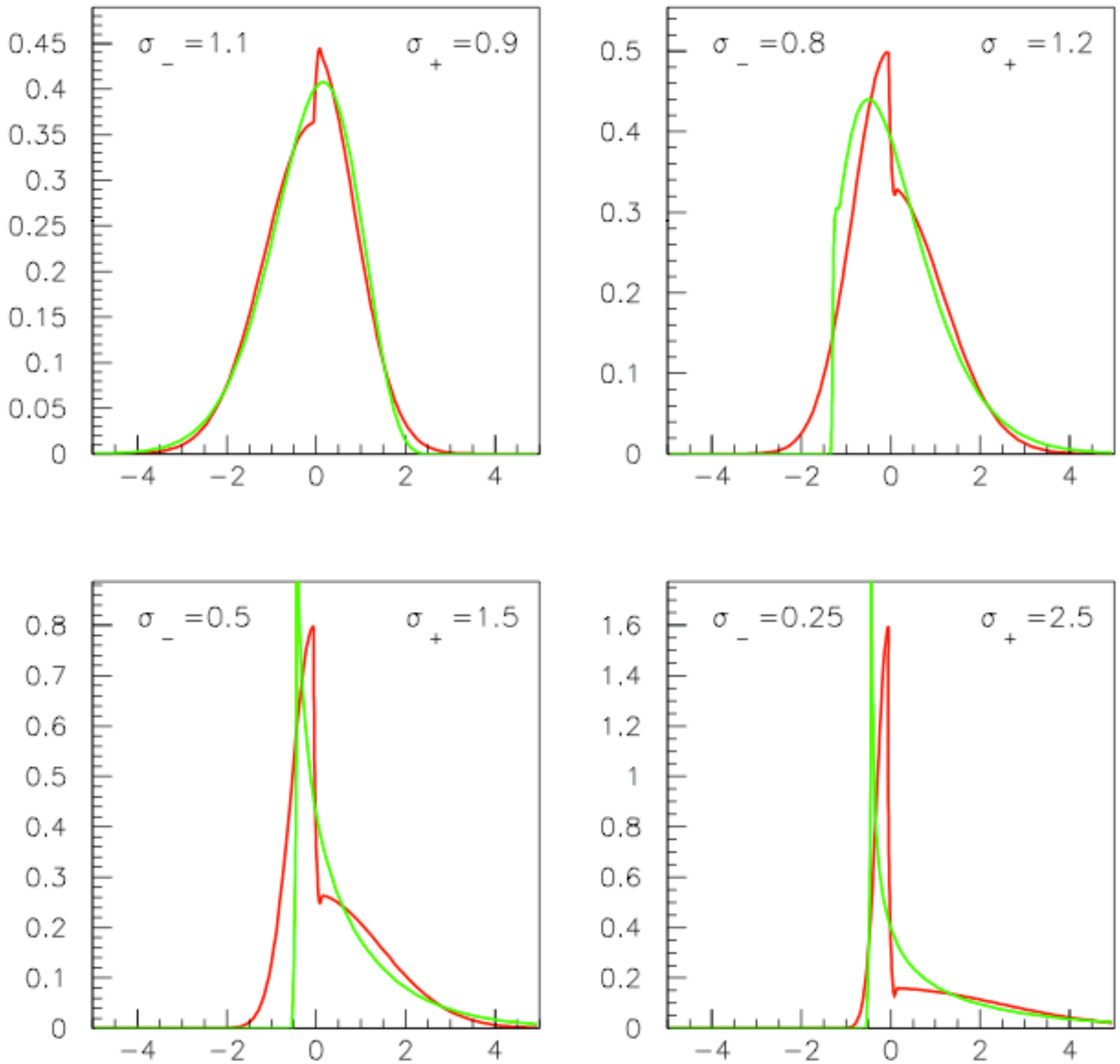
See R. Barlow,
"Asymmetric Systematic Errors", arXiv:physics/0306138
"Asymmetric Statistical Errors", arXiv:physics/0406120

# Quadratic Impacts of Asymmetric Uncertainties




R. Barlow

## Resulting Prior Distributions for alternative handling of Asymmetric Impacts




R. Barlow

# Even Bayesians have to be a little Frequentist

- A hard-core Bayesian would say that the results of an experiment should depend only on the data that are observed, and not on other possible data that were not observed.

  Also known as the "likelihood principle"

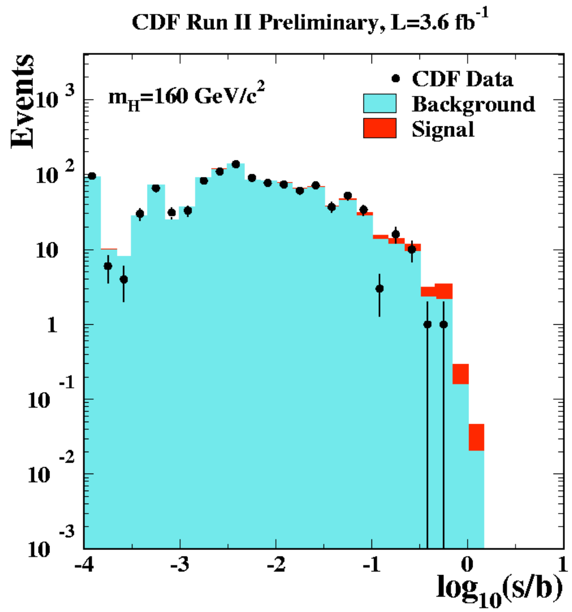
- But we still want the sensitivity estimated! An experiment can get a strong upper limit not because it was well designed, but because it was lucky.

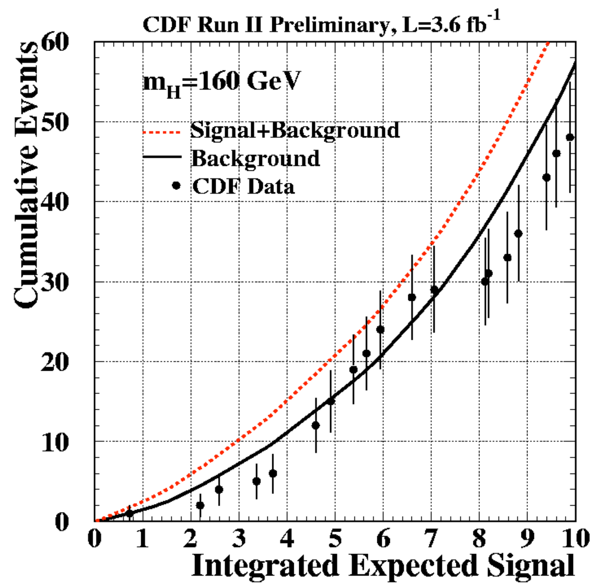
  How to optimize an analysis before data are observed?

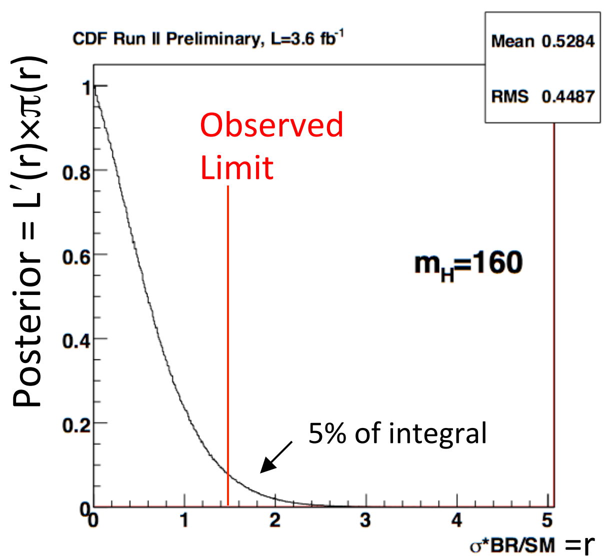
  So -- run Monte Carlo simulated experiments and compute a Frequentist distribution of possible limits. Take the median-- metric independent and less pulled by tails.

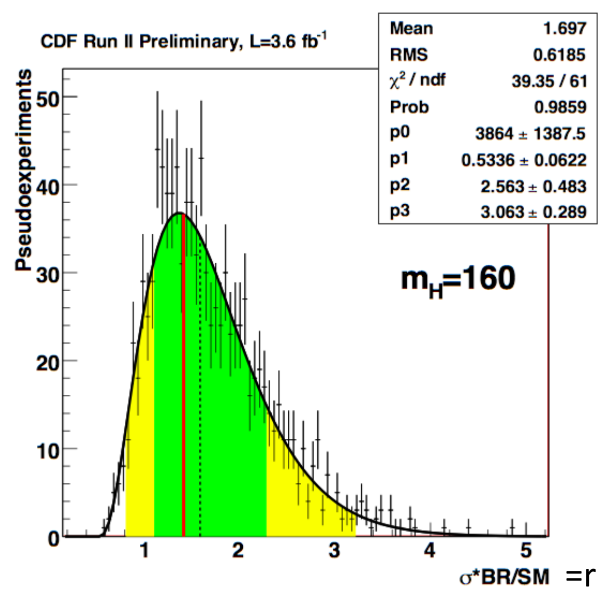
  But even Bayesian/Frequentists have to be Bayesian: use the Prior-Predictive method -- vary the systematics on eachc pseudoexperiment in calculating expected limits. To omit this step ignores an important part of their effects.

# Bayesian Example: CDF Higgs Search at $m_H$=160 GeV (an older one)

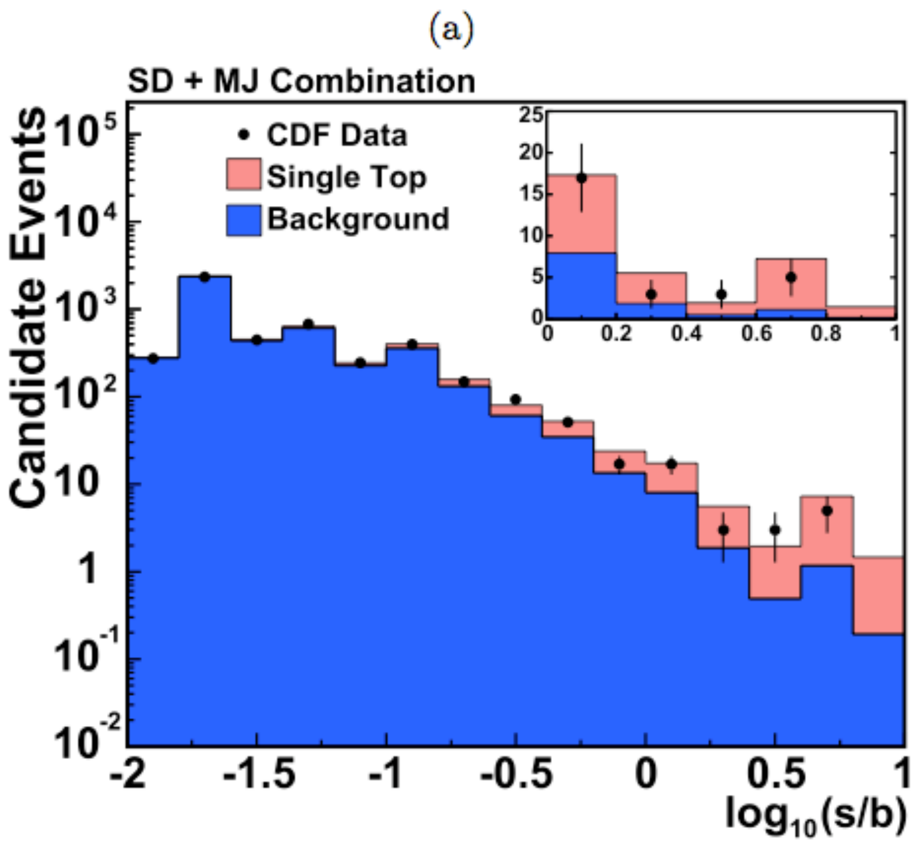


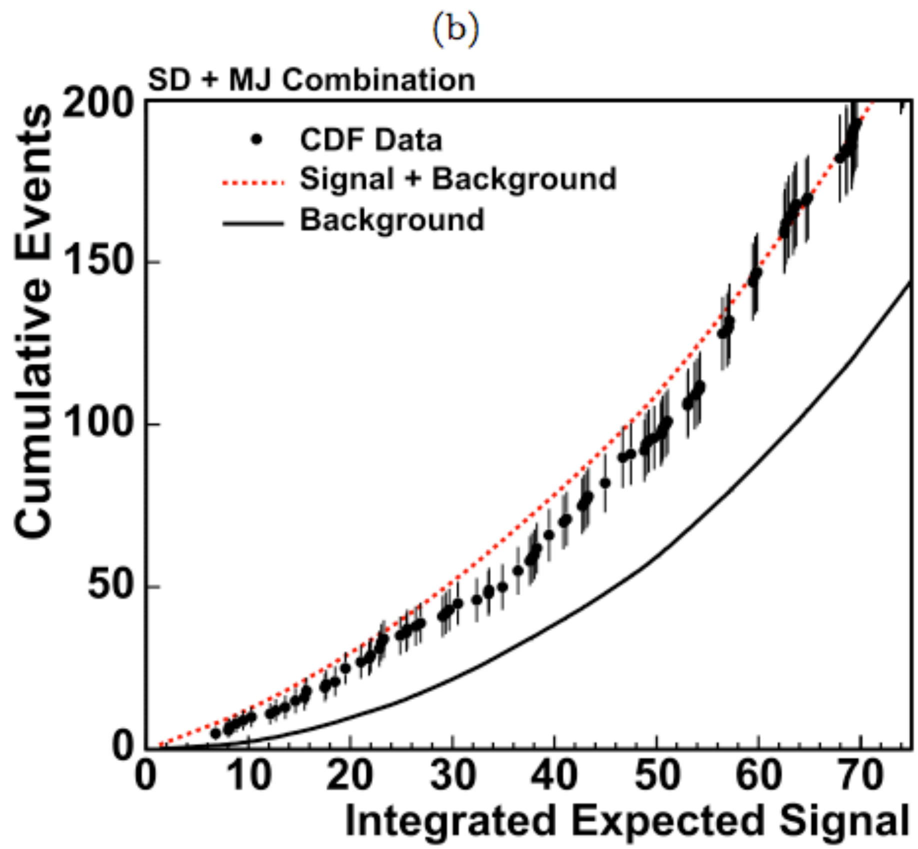



# What These Look Like for a 5.0σ Observation



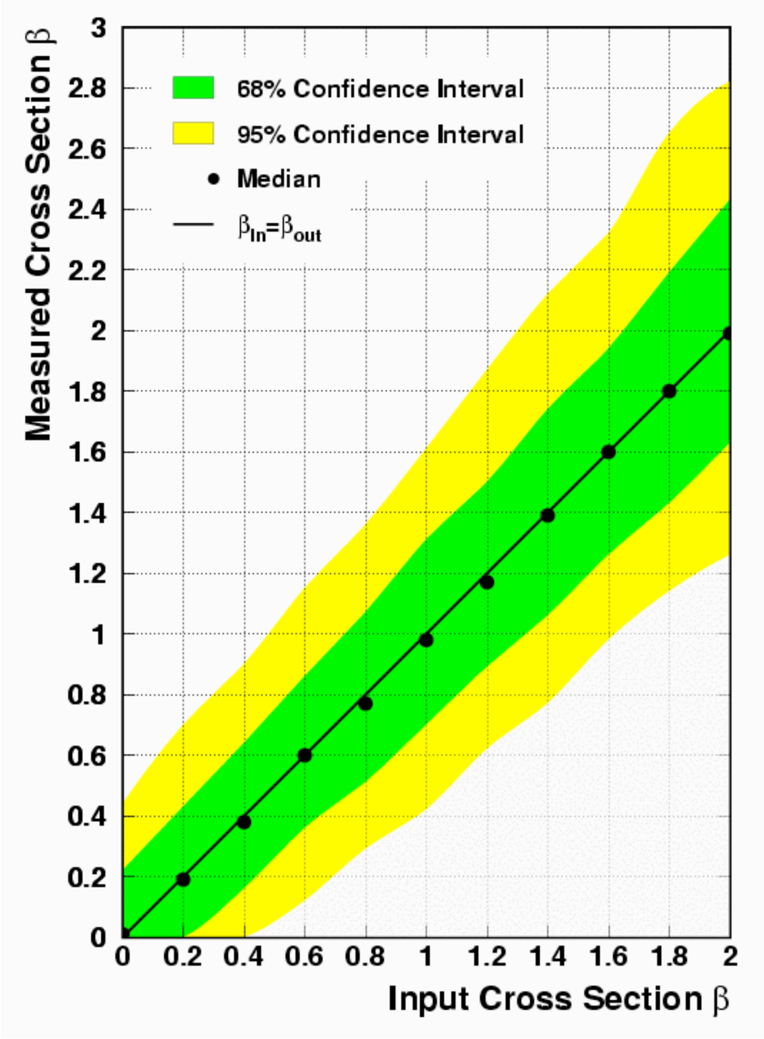

CDF Single Top, 3.2 $fb^{-1}$

# Even Bayesians have to be a little Frequentist

We would like to know how the cross section calculations behave in an ensemble of possible experimental outcomes.

Procedure:

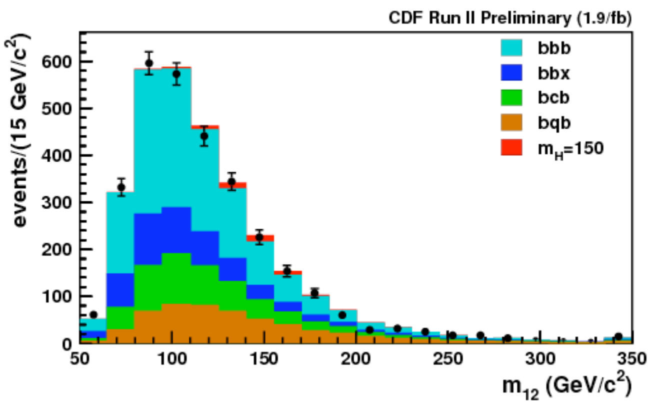
- Inject a signal.
- Vary systematics on each pseudoexperiment (which integrates over them in the ensemble)
- Calculate Bayesian cross section for each outcome and plot distribution.
- Black line is the median, not the mean
- Check the width of the distribution against the quoted uncertainties. Specifically, the distribution of (meas-inject)/uncertainty Should be a Unit-width Gaussian (when not up against zero).

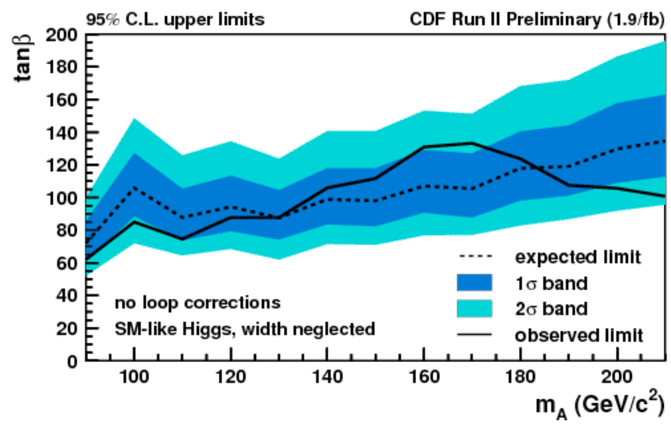


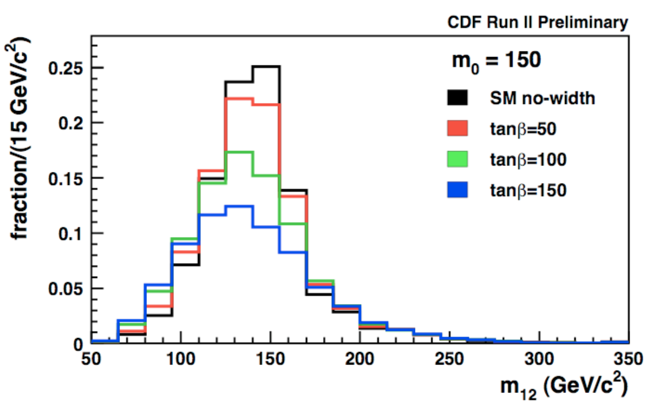

This is in fact a Neyman construction! Can do Feldman-Cousins with this (correct for fit biases, if any).

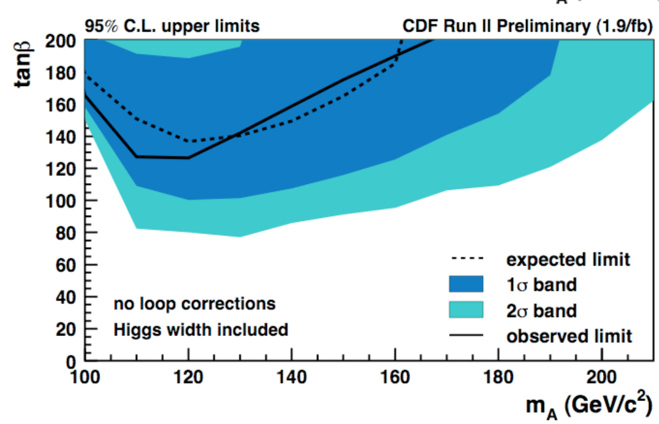
# An Example Where Usual Bayesian Software Doesn't Work

- Typical Bayesian code assumes fixed background, signal shapes (with systematics) -- scale signal with a scale factor and set the limit on the scale factor
- But what if the kinematics of the signal depend on the cross section? Example -- MSSM Higgs boson decay width scales with $\tan^2\beta$, as does the production cross section.
- Solution -- do a 2D scan and a two-hypothesis test at each $m_A$,$\tan\beta$ point



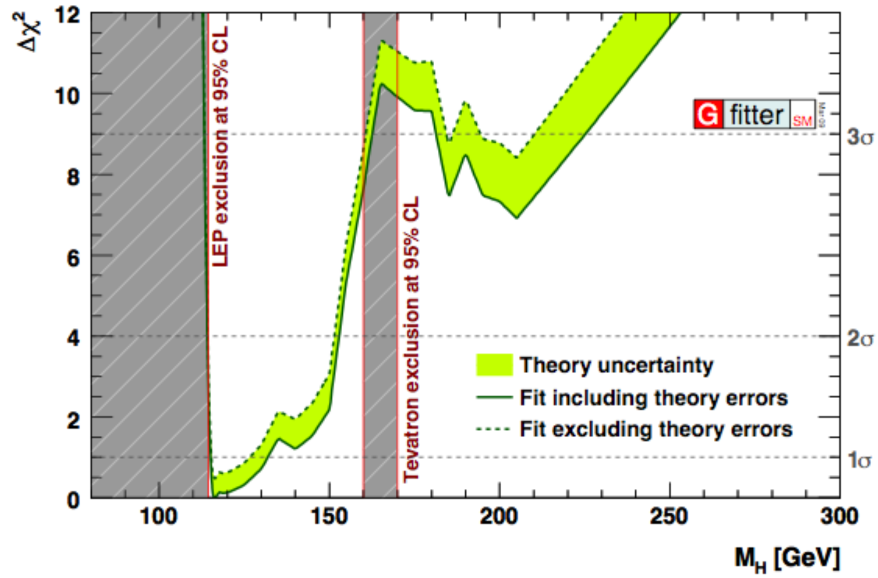



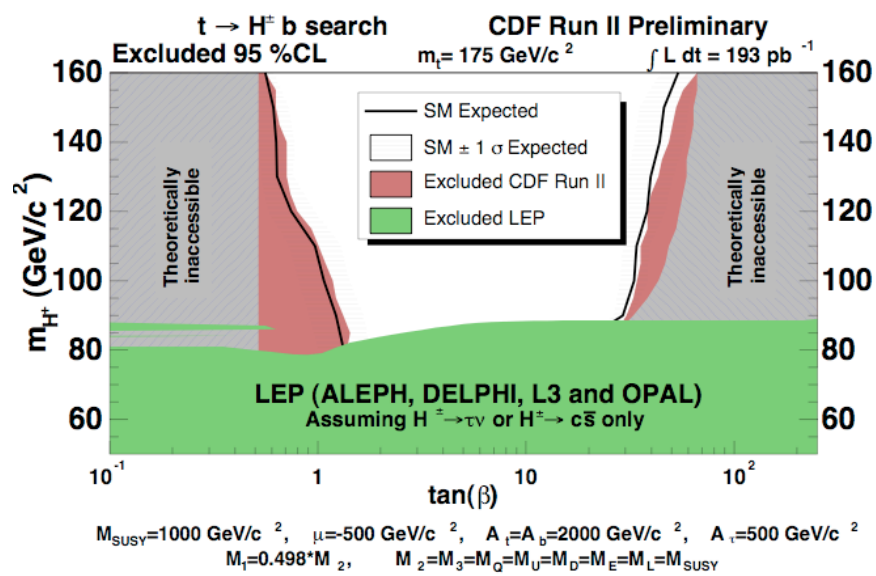



# Priors in Non-Cross-Section Parameters

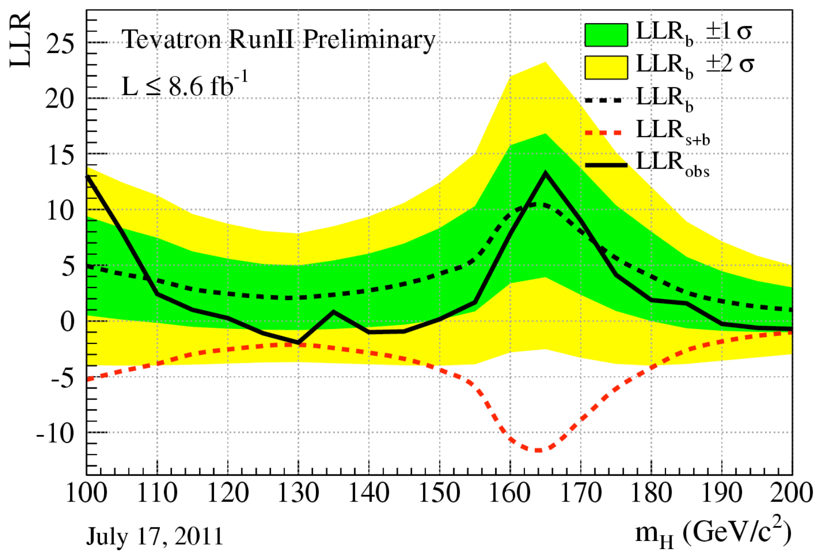
Example: take a flat prior in $m_H$; can we discover the Higgs boson by process of elimination? (assumes exactly one Higgs boson exists, and other SM assumptions)

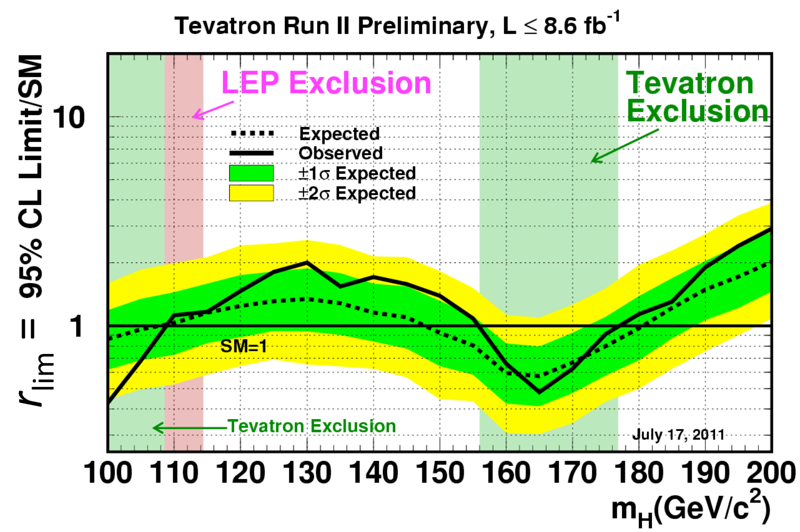


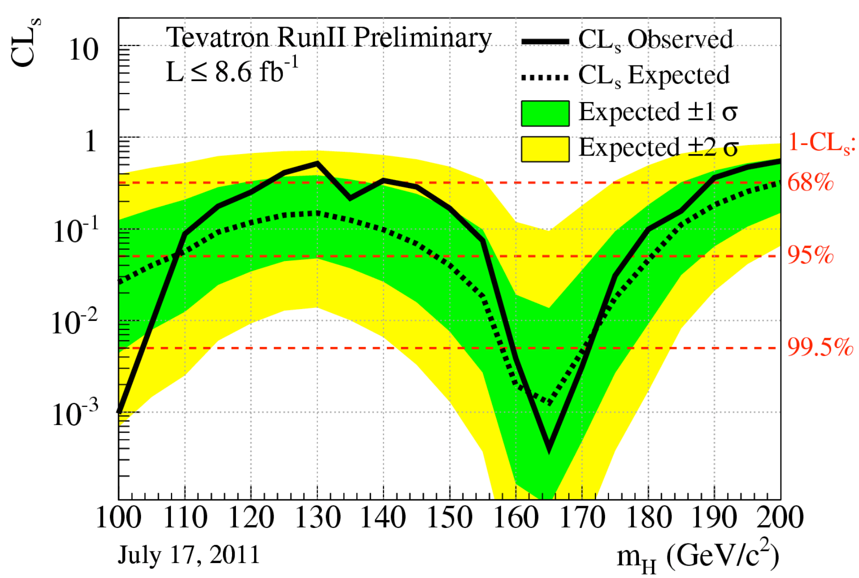

Example: Flat prior in log(tanβ) -- even with no sensitivity, can set non-trivial limits..



# Tevatron Higgs Combination Cross-Checked Two Ways




Very similar results --

- Comparable exclusion regions
- Same pattern of excess/deficit relative to expectation

n.b. Using $CL_{s+b}$ limits instead of $CL_s$ or Bayesian limits would extend the bottom of the yellow band to zero in the above plot, and the observed limit would fluctuate accordingly. We'd have to explain the 5% of $m_H$ values we randomly excluded without sufficient sensitivity.

# Measurement and Discovery are Very Different

Buzzwords:

- Measurement = "Point Estimation"
- Discovery = "Hypothesis Testing"

You can have a discovery and a poor measurement!
Example: Expected b=$2\times10^{-7}$ events, expected signal=1 event, observe 1 event, no systematics.

p-value ~$2\times10^{-7}$ is a discovery! (hard to explain that event with just the background model). But have ±100% uncertainty on the measured cross section!

In a one-bin search, all test statistics are equivalent. But add in a second bin, and the measured cross section becomes a poorer test statistic than the ratio of profile likelihoods.

In all practicality, discriminant distributions have a wide spectrum of s/b, even in the same histogram. But some good bins with b<1 event

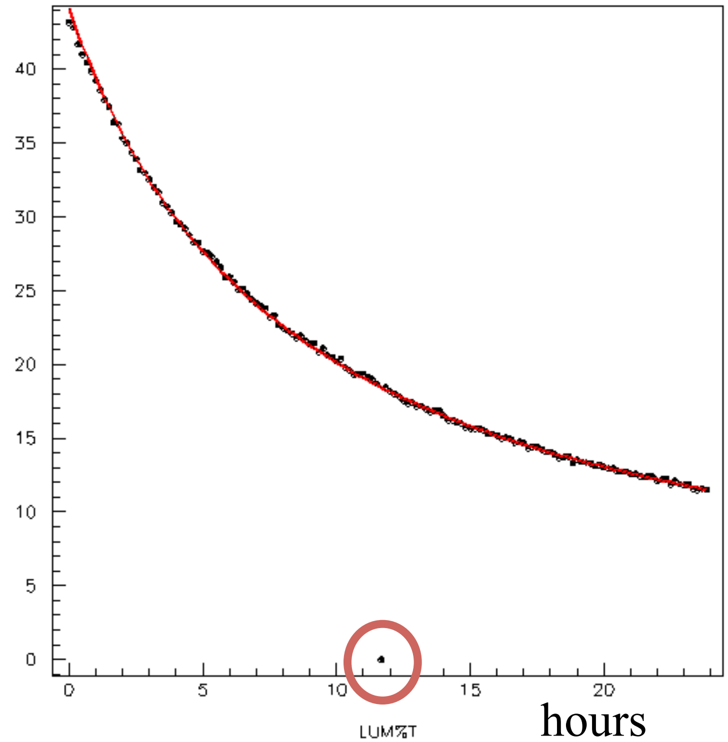
# Advantages and Disadvantages of Bayesian Inference

- Advantages:
  - Allows input of *a priori* knowledge:
    - positive cross-sections
    - positive masses
  - Gives you "reasonable" confidence intervals which don't conflict with *a priori* knowledge
  - Easy to produce cross-section limits
  - Depends only on observed data and not other possible data
  - No other way to treat uncertainty in model-derived parameters
- Disadvantages:
  - Allows input of *a priori* knowledge (AKA "prejudice") (be sure not to put it in twice...)
  - Results are metric-dependent (limit on cross section or coupling constant? -- square it to get cross section).
  - Coverage not guaranteed
  - Arbitrary edges of credibility interval (see freq. explanation)

# Outliers

- Sometimes they're obvious, often they are not.
- Best to make sure that the uncertainties on all points honestly include all known effects. Understand them!

Lum E30

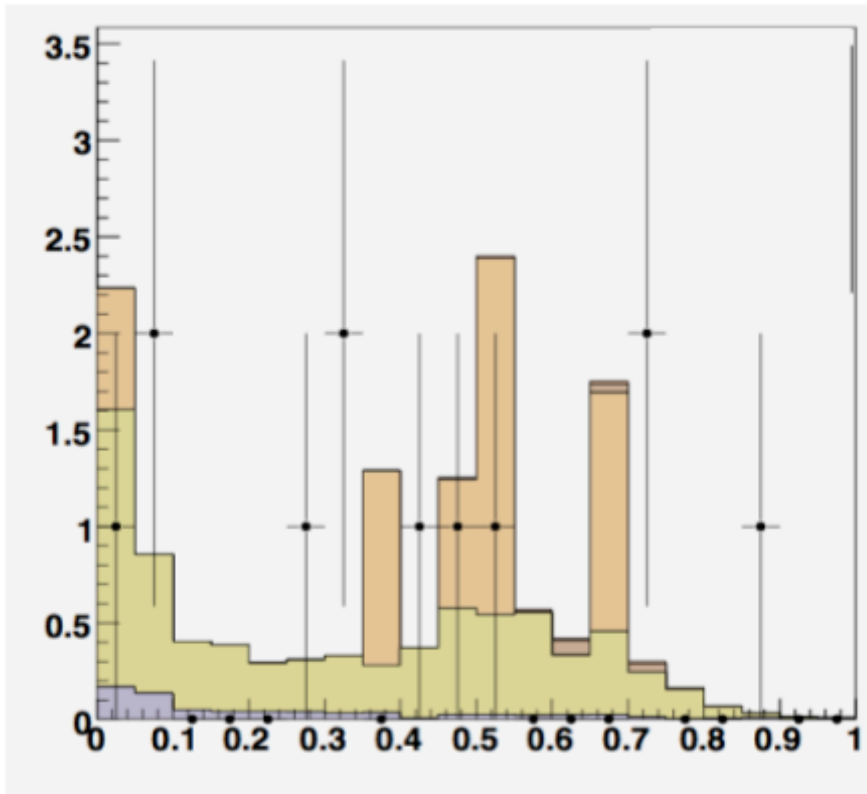



hours

L. Ristori, Instantaneous Luminosity at CDF vs. time (a Tevatron store in 2005)

# A Pitfall -- Not Enough MC (data) To Make Adequate Predictions

An Extreme Example (names removed)




Questions: What's the shape we are trying to estimate? What is the uncertainty on that shape?

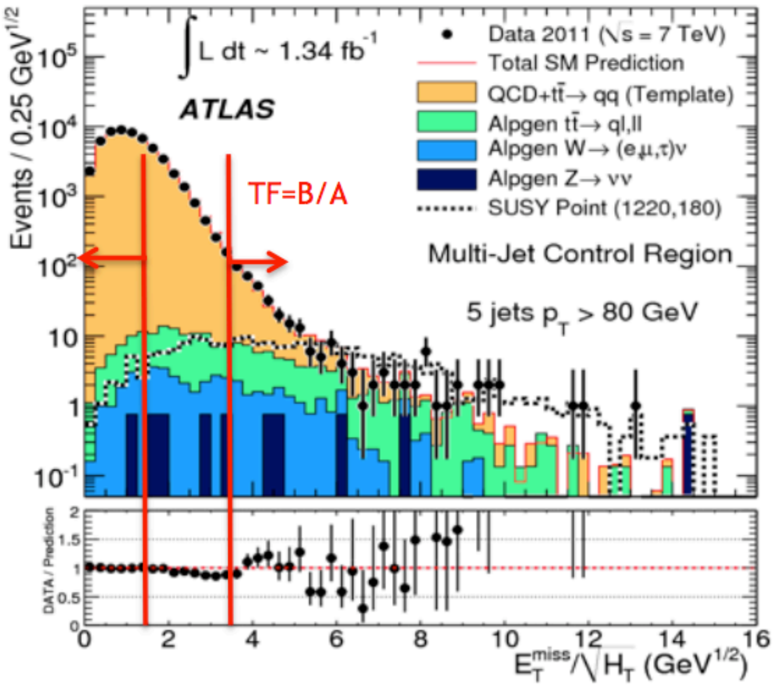
Cousins, Tucker and Linnemann tell us prior predictive p-values undercover with 0±0 events are predicted in a control sample.

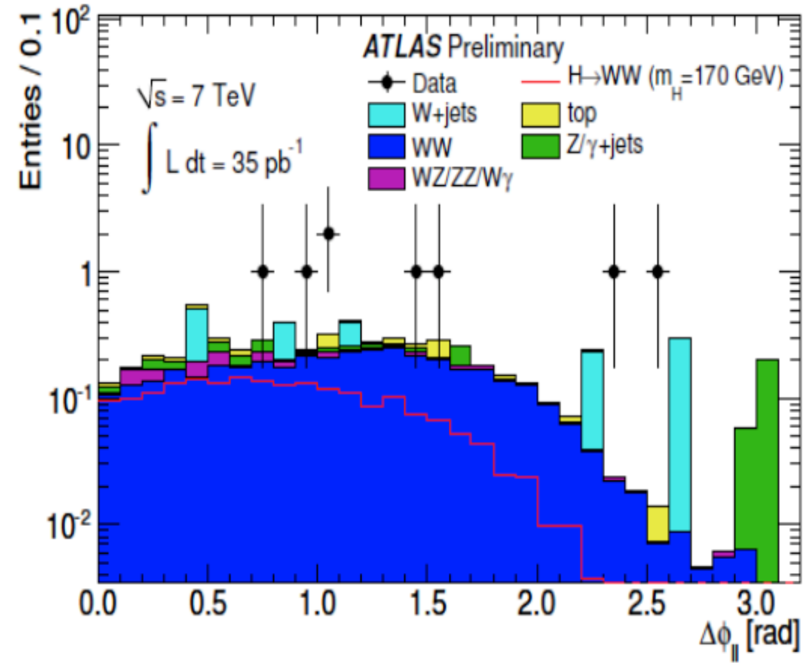
CTL Propose a flat prior in true rate, use joint LF in control and signal samples. Problem is, the mean expected event rate in the control sample is $n_{obs}+1$ in control sample. Fine binning → bias in background prediction.

Overcovers for discovery, undercovers for limits?

# Some Very Early Plots from ATLAS

Suffer from limited sample sizes in control samples and Monte Carlo
Nearly all experiments are guilty of this, especially in the early days!



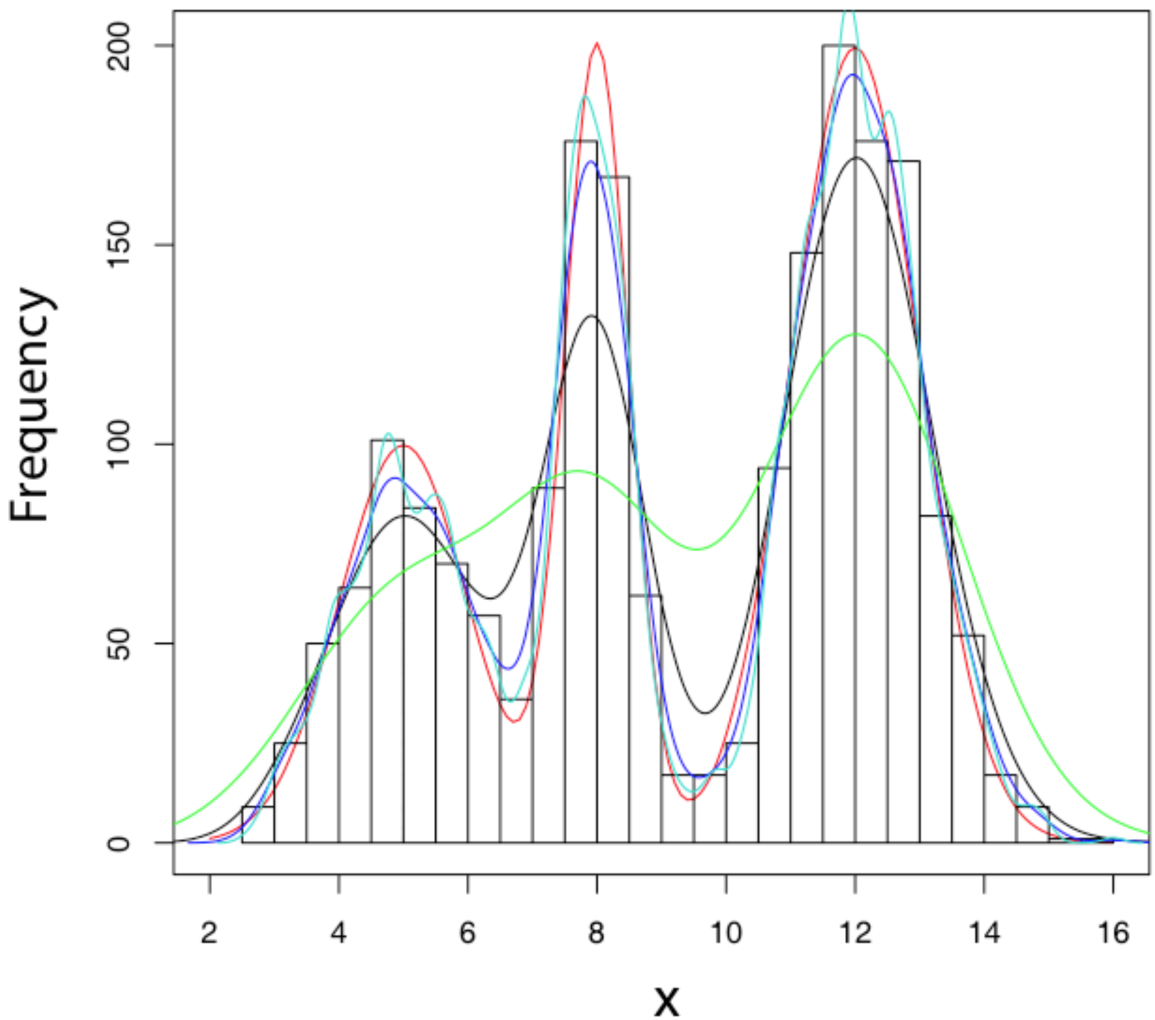

Data points' error bars are not sqrt(n). What are they? I don't know. How about the uncertainty on the prediction?

The left plot has adequate binning in the "uninteresting" region. Falls apart on the right-hand side, where the signal is expected.
Suggestions: More MC, Wider bins, transformation of the variable (e.g., take the logarithm).
Not sure what to do with the right-hand plot except get more modeling events.

# Smoothing Histograms

## Dependence on Smoothing Parameter

Plot showing effect of choice of smoothing parameter":




Red: Sampling PDF
Black: Default smoothing (w)
Blue: w/2 smoothing
Turquoise: w/4 smoothing
Green: 2w smoothing

Frank Porter, SLUO
lectures on statistics, 2006

# Optimizing Histogram Binning

Two competing effects:

1) Separation of events into classes with different s/b improves the sensitivity of a search or a measurement. Adding events in categories with low s/b to events in categories with higher s/b dilutes information and reduces sensitivity.

→ Pushes towards more bins

2) Insufficient Monte Carlo can cause some bins to be empty, or nearly so. This only has to be true for one high-weight contribution.

Need reliable predictions of signals and backgrounds in each bin

→ Pushes towards fewer bins

Note: It doesn't matter that there are bins with zero data events – there's always a Poisson probability for observing zero.

The problem is inadequate prediction. Zero background expectation and nonzero signal expectation is a discovery!

# Overbinning = Overlearning

A Common pitfall – Choosing selection criteria after seeing the data.
"Drawing small boxes around individual data events"

The same thing can happen with Monte Carlo Predictions –

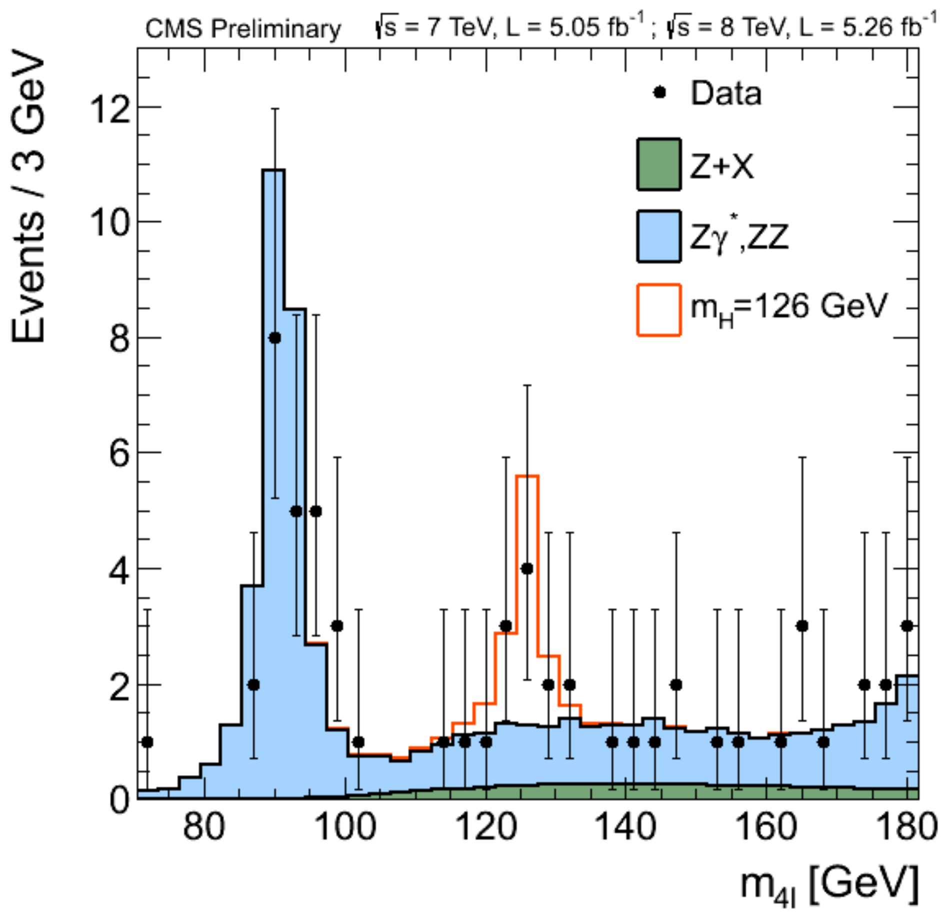
Limiting case – each event in signal and background MC gets its own bin.
→Fake Perfect separation of signal and background!.

Statistical tools shouldn't give a different answer if bins are shuffled/sorted.

Try sorting by s/b. And collect bins with similar s/b together. Can get arbitrarily good performance from an analysis just by overbinning it.

Note: Empty data bins are okay – just empty prediction is a problem. It is our job however to properly assign s/b to data events that we did get (and all possible ones).

# A Good Choice of Binning




CMS H→ZZ→4L

Bins with no data are fine!

Structures in signal and background are clear – not all bunched up into one bin

Sufficient signal and background predictions in each bin. We can interpret each event in the histogram by giving it a s/b

# A Comment on low s and low b

Bins with tiny s and tiny b can have large s/b (Louis Lyons: large s/sqrt(b) is suspicious)

Naturally occurring in HEP and others seeking discovery:

1) Each beam crossing has very small s and b but has the same s/b as neighboring beam crossings. Can make a histogram of the search for new physics separately for each beam crossing. Same s and b predictions, just scaled down very small.

   Adding is the same as a more elaborate combination if the histograms were accumulated under identical conditions (all rates, shapes, and systematics are the same)

2) Surveillance video catching a criminal – each frame has a small s, b, but still worthwhile to collect each frame (and analyze them separately)

# The 2011 CERN Unfolding/ Deconvolution Workshop

http://indico.cern.ch/conferenceOtherViews.py?view=standard&confId=107747

And look at the talks for Thursday, January 20, 2011 at the bottom of the page

# Available Software, Tools, Documentation

**CDF Statistics Committee**
http://www-cdf.fnal.gov/physics/statistics/statistics_home.html
Useful for documentation. Provides advice for common, thorny questions

**BaBar Statistics Working Group**
http://www.slac.stanford.edu/BFROOT/www/Statistics/

**ROOSTATS**
https://twiki.cern.ch/twiki/bin/view/RooStats/WebHome
A very complete toolset. I haven't used it (but I should have). It's in common use at the LHC

**MCLIMIT**
http://www-cdf.fnal.gov/~trj/mclimit/production/mclimit.html
Used on CDF, some use on D0 and LHC. Limits, cross sections, p-values, both Frequentist and Bayesian tools

**PHYSTAT.ORG**
http://www.phystat.org
Maintained by Jim Linnemann. We toolsmiths really should keep it up to date...

# Available Software, Tools, Documentation

**PDG Probability and Statistics Reviews (ed. Glen Cowan)**

http://pdg.lbl.gov/2012/reviews/rpp2012-rev-probability.pdf
http://pdg.lbl.gov/2012/reviews/rpp2012-rev-statistics.pdf

If these links get out of date, just search pdg.lbl.gov for the mathematical reviews
Excellent brief reference, but maybe a little too brief to learn the material.

**Good Reads:**

Frederick James, "Statistical Methods in Experimental Physics", $2^{nd}$ edition, World Scientific, 2006

Louis Lyons, "Statistics for Nuclear and Particle Physicists" Cambridge U. Press, 1989

Glen Cowan, "Statistical Data Analysis" Oxford Science Publishing, 1998

Roger Barlow, "Statistics, A guide to the Use of Statistical Methods in the Physical Sciences", (Manchester Physics Series) 2008.

Bob Cousins, "Why Isn't Every Physicist a Bayesian" Am. J. Phys **63**, 398 (1995).

# Extra Material

# Analysis Optimization in Isolation or in Combination?

Typical situation:

A measurement has a statistical and a systematic uncertainty, where the statistical uncertainty includes "good" systematics that are constrained by the data, and the "bad" ones never get better constrained no matter how much data are collected.

We sometimes have a choice of how to analyze marked Poisson data.

1) aggressive reconstruction making assumptions about particle distributions – more statistical power per event at the cost of introducing systematic uncertainty

2) more model-independent analysis with fewer assumptions – less statistical power per event but better control over systematics.

→ Combination with other measurements (from other data runs or other collaborations) is like collecting more data. Method 1 hits the systematic limit and loses weight in the combination even though it may be the most powerful method by itself.

More general: With little data, we are more dependent on our assumptions, with more data we can relax the assumptions and constrain our models.

Recommendation: For combinations, optimize for the large luminosity case.

# Statistical Uncertainties on Systematic Uncertainties?

Answer: No. But some systematic uncertainties are difficult to evaluate properly.

See Roger Barlow's "Systematic errors: Facts and Fictions",
arXiv: hep-ex/0207026

The idea: If a systematic uncertainty is estimated by comparing two data samples or two MC samples, or data vs. MC, then if one or both of them have a limited size, then the magnitude of the systematic can be poorly constrained.

Ideally, work harder (run more MC) to get a better prediction of the expected signal and background, under all assumptions of systematic variation.

**Monte Carlo Statistical Uncertainty is a Systematic Uncertainty**

but don't double-count it for each separate MC variation of each nuisance parameter. Easy to do by comparing central vs. varied MC samples.

Statistically weak tests should be handed as cross checks. If they are consistent, consider the test to have passed, but do not add systematic uncertainty.
If they fail, however, and a discrepancy between two MC's or data and MC cannot be understood and fixed, then a systematic uncertainty is called for.

# Bayesian Discovery?

Bayes Factor

$$B = L'(data \mid r_{\max}) / L'(data \mid r = 0)$$

Similar definition to the profile likelihood ratio, but instead of maximizing L, it is averaged over nuisance parameters in the numerator and denominator.

Similar criteria for evidence, discovery as profile likelihood.

Physicists would like to check the false discovery rate,
and then we're back to p-values.

But -- odd behavior of *B* compared with p-value for even a simple case

J. Heinrich, CDF 9678
http://newton.hep.upenn.edu/~heinrich/bfexample.pdf

# Correlations among Uncertainties – When is it Conservative, when not?

- Within a channel – contributions that add together: including correlations usually weakens the sensitivity (always: sensitivity is expected)
- Between channels – accounting for correlations is not conservative
  One channel's observed data becomes another "off" sample for another's.
  Have to trust all the τ factors, and even offsets from central predictions
  in order to put in these correlations.

- Overestimating the impacts of systematic uncertainty on a prediction is not conservative if a correlation is taken into account. Can result in underestimated systematic error on a combined result.

Example (systematic uncertainty 1 is 100% correlated, syst uncertainty 2 is 100% correlated

Measurement 1: $m_1 = 5 \pm 1$ (syst1) $\pm 1$ (syst2) Combine with BLUE: $m_{best}=2m_1-m_2$
Measurement 2: $m_2 = 5 \pm 1$ (syst1) $\pm 2$ (syst2) → $m_{best}=5 \pm 1$ (syst1) $\pm 0$ (syst2)

Here accounting for correlation and an overestimated systematic uncertainty results in an aggressive result.

# Binned and Unbinned Analyses

- Binning events into histograms is necessarily a lossy procedure

- If we knew the distributions from which the events are drawn (for signal and background), we could construct likelihoods for the data sample without resort to binning. (Example Next page)

- Modeling issues: We have to make sure our parameterized shape is the right one or the uncertainty on it covers the right one at the stated C.L.

- Unfortunately there is no accepted unbinned goodness-of-fit test

A naive prescription: Let's compute L(data|prediction), and see where it falls on a distribution of possible outcomes –
compute the p-value for the likelihood.

Why this doesn't work: Suppose we expect a uniform distribution of events in some variable. Detector ϕ is a good variable. All outcomes have the same joint likelihood, even those for which all the data pile up at a specific value of phi. Chisquared catches this case much better.

Another example: Suppose we are measuring the lifetime of a particle, and we expect an exponential distribution of reconstructed times with no background contribution. The most likely

# Sensitivity of upper limit to Even a “flat” Prior

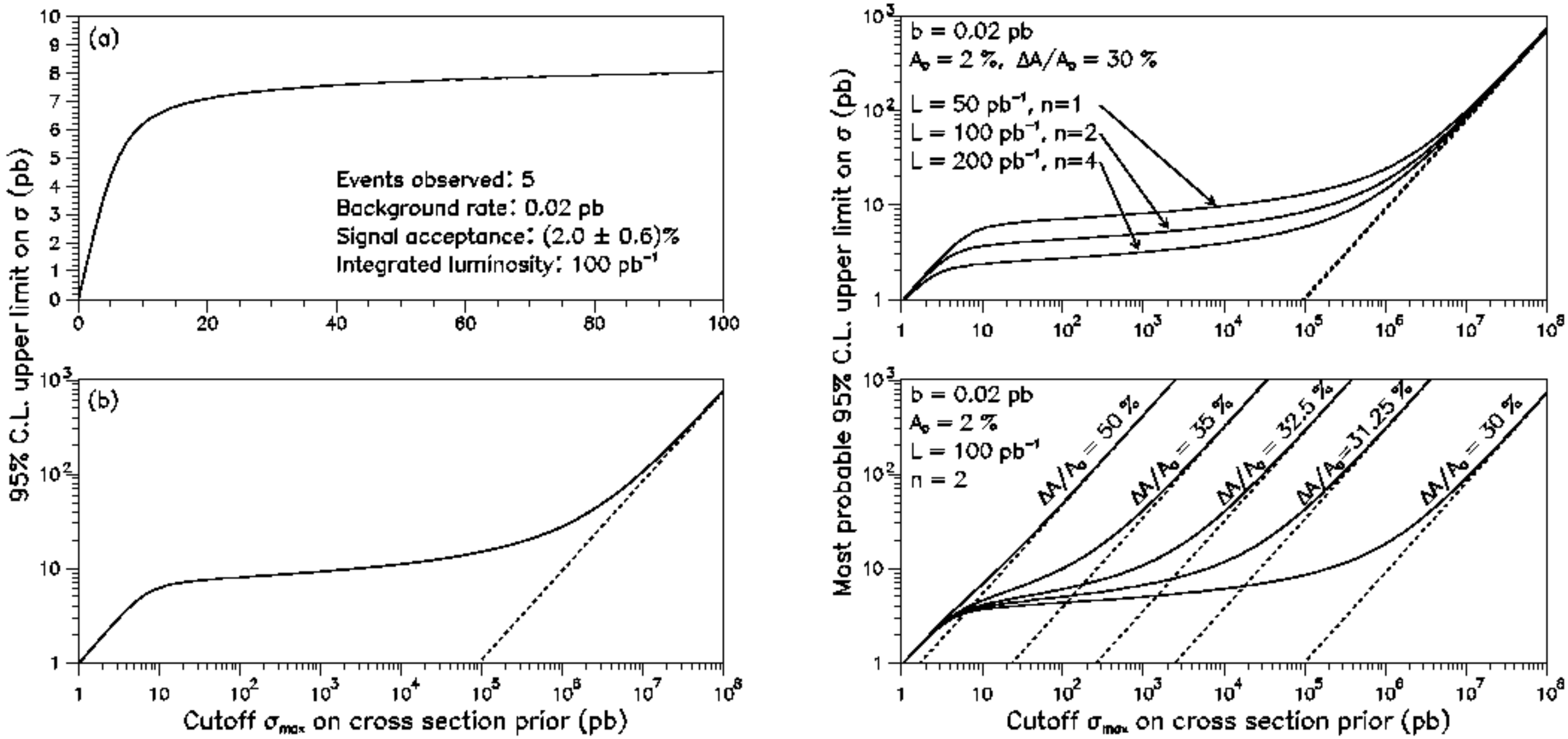



Figure 1: Bayesian upper limits at the 95% credibility level on a hypothetical cross section $\sigma$, as a function of the cutoff $\sigma_{\max}$ on the flat prior for $\sigma$.

L. Demortier, Feb. 4, 2005

# Example of a Pitfall in Fitting Models

- Fitting a polynomial with too high a degree
- Can extrapolations be trusted?

Trigger x-section extrapolation vs. luminosity

CEM16_TRK8