

# ***Limits to Discovery: Extended discussion***

- x Practical Issues
- x Ideological Issues

**This short intro is intended to provoke discussion and is by no means a thorough review**

Wade Fisher  
*Fermilab*

***UK HEP Forum Discussion Session***

***May 7<sup>th</sup> 2009***

Goal: Evaluate a statistical result using some number of input channels.

What ingredients will impact your result?

For authors of statistical software:

What choices must you make in order to deliver a robust product?

For the teams who combine input channels:

What are the show-stoppers? Where are the speed-bumps?

For those who will interpret the quality of the result:

What assumptions were made? How do they impact the result?

- x Assuming you've already made the philosophical choices of paradigm, test statistic, etc.

Where are the sharp edges upon which users may impale themselves?

- x How do you handle overflow bins?

Analyzers frequently do not consider the contents of overflow bins.

This leads to unexpected (and often impossible to find) errors.

Suggestion: Ignore overflow bins. Analyzers can add these to the histogram if they are serious about it.

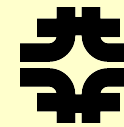
- x How do you handle statistical uncertainties?

Bin-by-bin statistical uncertainties can be the largest uncertainty in a calculation.

Did the analyzer get this right? How do you know?

Suggestion: Integrate careful tests of bin-by-bin errors and reject them (or flag them) if they are suspect. (error larger than bin content, error==0, etc)

# For Software Authors



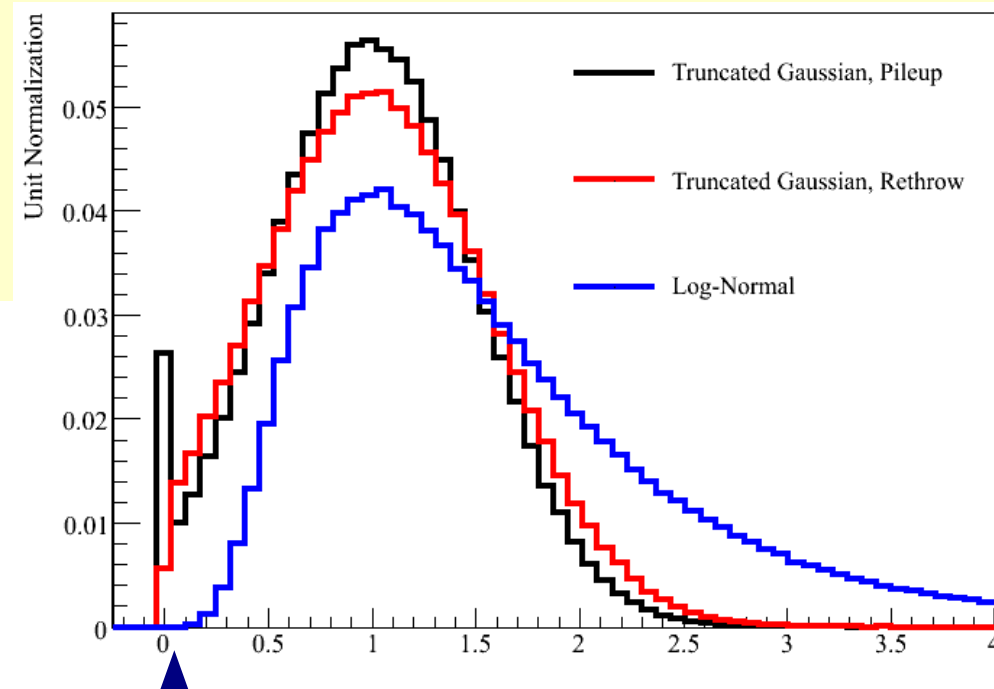
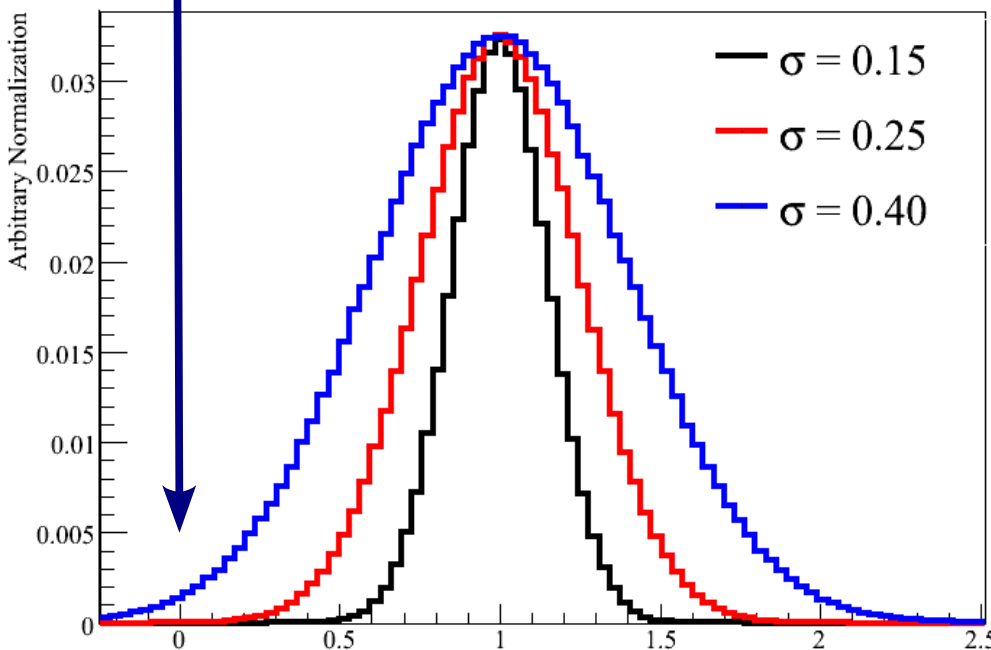
UK HEP  
May 7<sup>th</sup>, 2008

✗ How does one model systematics which vary as much as 50%?

We generally model uncertainties with Gaussian priors.

Lesson: Over-inflated systematics have implications beyond the obvious ones

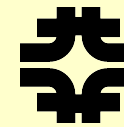
Non-physical, negative values  
populated as  $\sigma$  grows.



Must choose appropriate model  
for uncertainties approaching  
zero rate.

- x Given a list of input channels, what stands between you and your results?
  - x What is your deadline?
    - Do you have enough time? Will the inputs change mid-calculation?
    - Suggestion: Be very pragmatic about deadlines and the freezing of inputs.  
Don't set yourself up to fail.
- x Are you prepared to combine channels (or experiments)?
  - Are all physics model choices identical? Are systematic correlated?
  - If so, do all inputs use the same assumptions and conventions?
  - Suggestion: The combination effort begins many months before the actual combination. Decide on combination parameters at analysis design time.
- x Analyzers are harried! How can you alleviate the burden on analyzers?
  - Relative distribution of effort in bug fixes, upgrades, mods etc is important.
  - Suggestion: Implement an input format framework that is as flexible as possible. Educate analyzers on conventions early on.

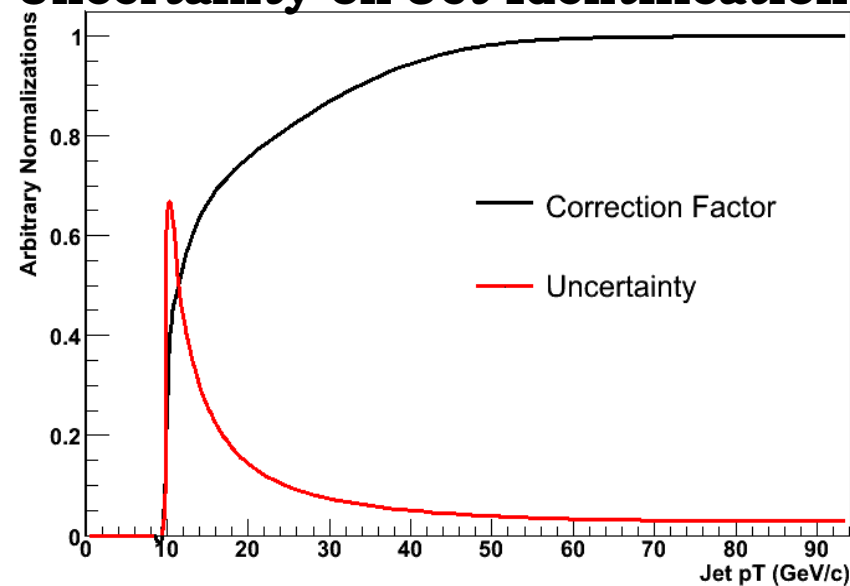
# For Reviewers



UK HEP  
May 7<sup>th</sup>, 2008

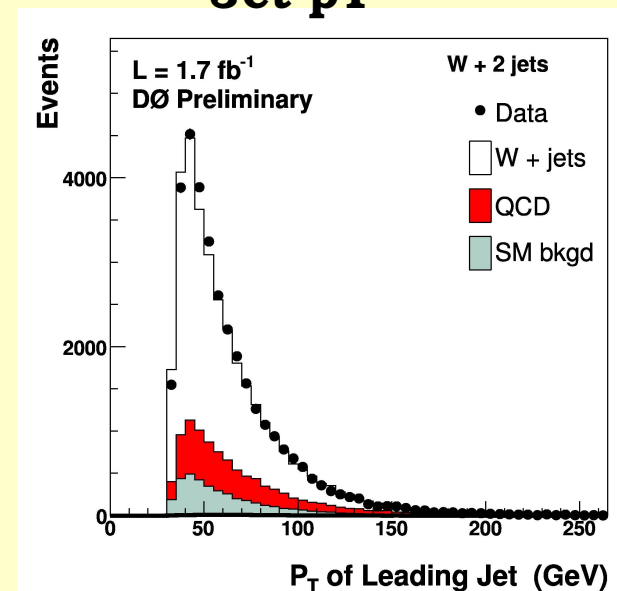
- × Presented with a result, how can you judge the quality?  
What assumptions are likely to have been made?
- × Gaussian truncation  
See earlier slide....
- × What assumptions were made about nuisance parameter priors?  
We generally assume Gaussian and parametrize by  $1\sigma$ . Does  $3\sigma = 3 \times 1\sigma$ ?  
Must consider impact on input, not nuisance parameter.

## Uncertainty on Jet Identification

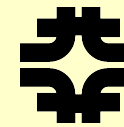


6

## Jet pT



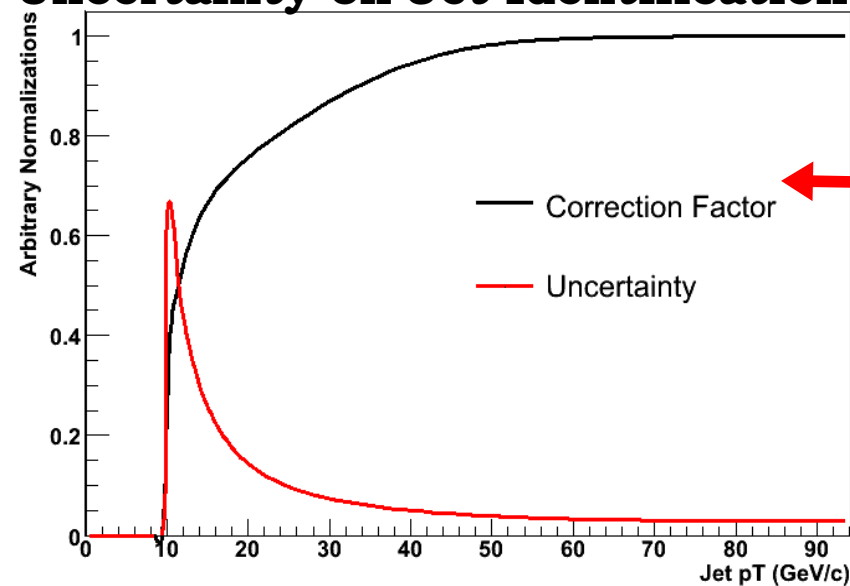
# For Reviewers



UK HEP  
May 7<sup>th</sup>, 2008

- × Presented with a result, how can you judge the quality?  
What assumptions are likely to have been made?
- × Gaussian truncation  
See earlier slide....
- × What assumptions were made about nuisance parameter priors?  
We generally assume Gaussian and parametrize by  $1\sigma$ . Does  $3\sigma = 3 \times 1\sigma$ ?  
Must consider impact on input, not nuisance parameter.

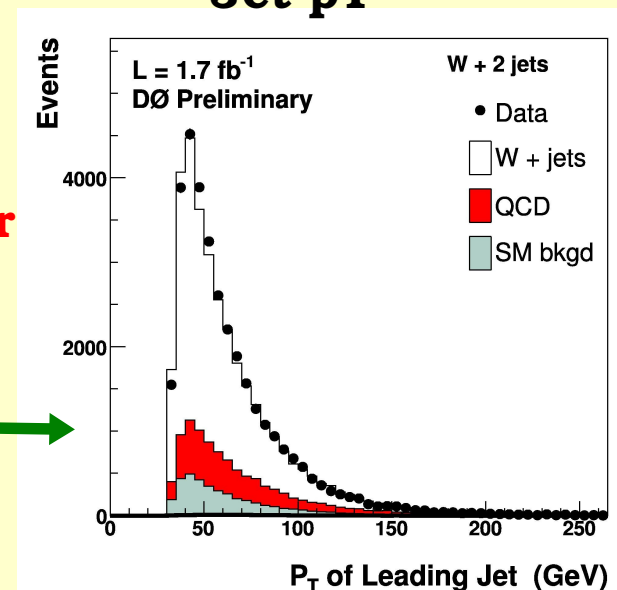
## Uncertainty on Jet Identification



Convolute  $1\sigma$  on  
nuisance parameter  
with the  
corresponding  
observable

7

## Jet pT

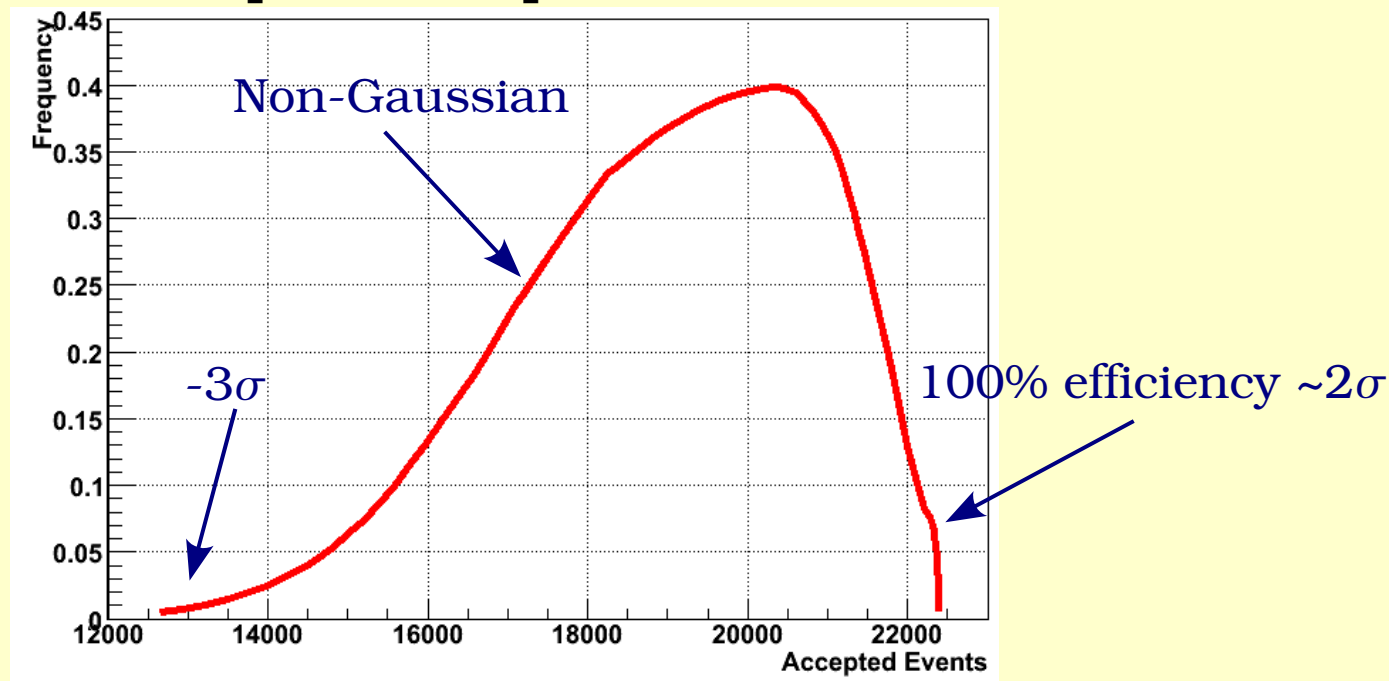


- × Presented with a result, how can you judge the quality?  
What assumptions are like to have been made?
- × What assumptions were made about nuisance parameter priors?  
We generally assume Gaussian and parametrize by  $1\sigma$ . Does  $3\sigma = 3 \times 1\sigma$ ?  
Must consider impact on input, not nuisance parameter.

**Potentially large  
impact on discovery  
calculations.**

**For 5-sigma  
observation, are you  
frequently probing  
 $3\sigma$  for any  
systematics?**

### Acceptance Response Function





Outstanding questions and issues to consider.

What questions will your colleagues ask of you? Your competitors?

x What is the meaning and benefit of cross-checks with multiple techniques?

Disagreement at which level identifies a problem?

In which cases should different techniques give different answers?

x Coverage:

Do we care about coverage only in the parameter of interest?

What about nuisance parameters?

Bayes only or profile likelihood as well?

x Priors:

Is a non-informative prior the best (only) way to go with Bayes?

Do we learn anything from clever priors? How sensitive are we to the choice?

x Outstanding questions and issues to consider.

What questions will your colleagues (competitors?) ask of you?

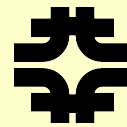
x Exclusion vs. Discovery

Do we require a consistent (identical?) technique or do we cherry pick?

Do we choose the best or worst result?

Should this depend on the problem?

# Summary



UK HEP  
May 7<sup>th</sup>, 2008

- × A summary list of topics that tend to plague the folks in the front lines of statistical calculations
  - A few nitty-gritty practical issues
  - A few more philosophical issues
  - Each area is equally important to producing a quality result
  
- × Mostly questions & suggestions, no manna from heaven
  - Hopefully this is a good start for discussion
  - Not a thorough dissertation on the topics