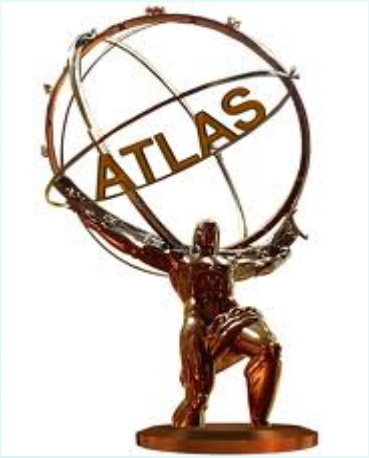
Strategy

6th December 2010

Bill Murray
LHC-HCG



- Software framework
- Statistical methods
- Time-lines for work

Combination Software

- The assumed frame is ROOSTATS.

  https://twiki.cern.ch/twiki/bin/view/RooStats/WebHome
  - This has been agreed by the stats. fora for years and developed largely for this job.
    - However there are questions about speed and reliability.
  - We intend to use RooStats workspaces to export and import the channels or experiment codes.
  - Functional forms versus histograms needs to be clarified.
    - Histograms are clearly acceptable;
    - we need to consider functions carefully before adopting.
- We need advice from stat. forum reps...

Validation

- Essential to repeat and extend the validations of RooStats against others
  - e.g. MCLimit, MatLab
  - various private codes which have been developed..)
  - on test examples but primarily on the H→WW combination.
- As the inputs will be produced in RooStats form, any others codes will need to interface to those.

Limit Extraction

- One advantage of the ROOSTATS frame is the flexibility of choosing the statistical method.
- **1** will be adopted, comparison with ?1? other
  - To Illustrates the influence of method
  - And the meaning of the statistic
- A plausible choice, we can discuss, would be
  - CL_s or Power Constrained Limits
  - with a Bayesian with Jeffries prior comparison.
- We do not intend to allow this to become a major issue. It is an (important) technical detail.
  - A decision is required BEFORE combination

Discovery

- Discovery through p-value significance,
  - quoted both at the mass concerned,
  - and allowing for look-elsewhere.
    - But the window range is always contentious – so this is informational
- Using the Profile Likelihood
- Will study Asimov(typical) v toy MC approaches
  - Toy MC is accurate but expensive
  - Asymptotic formulae, (related to χ^2) inaccurate if n too low
  - Hopefully both agree in reasonable range

Common Systematics

- Understanding common systematics will be an important issue, especially in cross-sections.
- The representatives of the cross-section working group will be required to bring expertise.
  - e.g. * For different acceptance cuts, what is the correlated/uncorrelated signal error?
  - For similar control regions used to estimate backgrounds, how correlated is the result.
- Luminosity: Currently 100% correlated, but...
- Need to allow time for these studies
  - May need pragmatic solutions

When does group need inputs?

- Methods 8 weeks before
  - group members know what analyses each experiment plans
  - this must be kept confidential to this group
    - ATLAS/CMS should not know each other's plans
- Data 2-3 weeks before
  - Even more confidential, until....
  - Results to collaborations a week before meeting

First steps

- WW channel repeat,
  - Studying systematics in details
  - Keep production mechanism contributions identifiable if possible
- Build up from 1 channel
  - Adding different WW channels (0/1/2 jets)
  - Maybe other WW decay modes?

Combination preparation

- Check sensitivities of channels in comparisons between experiments.
- Understand big differences/discrepancies.
  - i.e. explain differences in terms of selections / detector performance..
  - Do we need to have results believable by both collaborations?
- Recheck stat. technique
  - Accuracy of asymptotic formulae for significance
  - Evaluate speed of combination code

Final Results preparation

- Important to reproduce each others results
- If we use a Roostats with a bug, we will all get same buggy result..
  - Do we need to reproduce results with different codes AT THIS POINT?
    - i.e. In addition to previous validation
- Combination of data should be quick as we have prepared beforehand
- But mistakes **will** happen and we must catch as many as possible.

Conclusions

- This isn't going to be easy
- We need to trust each other
  - Work fast
    - Will need small ad-hoc meetings
  - But independently until close to the end
- People are expecting these results fast.
- We must start working on this now