TOWARDS A WHITE-PAPER ON PUBLIC LIKELIHOODS

@KyleCranmer
New York University
Department of Physics
Center for Data Science
CILVR Lab
I have a lot of thoughts on this topic, so aiming at high-level points

- Please see PhyStat seminar for more details that may be helpful for the white paper https://indico.cern.ch/event/962997/

See also Nicholas Wardle’s PhyStat Seminar:

- https://indico.cern.ch/event/1012319/
Introduction

Topics:

• Technical issues around likelihoods
• Surrounding (cyber)infrastructure: repositories, citation, RECAST, …
• Incentives, encouragement, evidence to support those in community working to make this happen
• Goals of white paper, audience
We are at a “tipping point” There is more positive momentum than I’ve ever seen

• Suggest we take a positive tone of encouragement

• We should also all educate ourselves about all the efforts that are going on:
  • Citations to DOIs for software and data products
  • Training efforts

My sense is that within different groups (Higgs, SM, SUSY/Exotics, top, flavor) and different experiments (ATLAS, CMS, LHCb, …) there are different attitudes

• e.g. ATLAS SUSY/Exotics embracing full likelihoods, opening discussion more broadly

• but surprisingly Higgs EFT and STXS community don’t seem to be very aware of developments or relevance to their problems
REPRODUCIBILITY PROBLEM

Not possible for others to reproduce results from paper.

Figure 4: Fits for 2-parameter benchmark models probing different coupling strength scale factors for fermions and vector bosons: (a) Correlation of the coupling scale factor $\kappa_F$ and $\kappa_V$, assuming no non-SM contribution to the total width; (b) Correlation of the coupling scale factors $\lambda_{FV} = \kappa_F / \kappa_V$ and $\kappa_{VV} = \kappa_V \cdot \kappa_V / \kappa_H$ without assumptions on the total width.

Figure 5: Fits for benchmark models probing different coupling strength scale factor for fermions and vector bosons, assuming no non-SM contribution to the total width: (a) coupling scale factor for fermions $\kappa_F$ (the coupling scale factor for gauge bosons $\kappa_V$ is profiled) and (b) coupling scale factor for gauge bosons $\kappa_V$ (the coupling scale factor for fermions $\kappa_F$ is profiled).
Not possible for others to reproduce results from paper.
**LIKELIHOOD SCANS**

**First step:** publish likelihood scans for communicating LHC Higgs results.

http://doi.org/10.7484/INSPIREHEP.DATA.A78C.HK44

http://doi.org/10.7484/INSPIREHEP.DATA.RF5P.6M3K

http://doi.org/10.7484/INSPIREHEP.DATA.26B4.TY5F

---

**Measurements of Higgs boson production and couplings in diboson final states with the ATLAS detector at the LHC**

ATLAS Collaboration (Georges Aad (Freiburg U.) *et al.*) Show all 2923 authors

Jul 4, 2013 - 32 pages


---

Data from Figure 7 from: Measurements of Higgs boson production and couplings in diboson final states with the ATLAS detector at the LHC

ATLAS Collaboration (Aad, Georges (Freiburg U.) *[]) Show all 2923 authors

Reproducing derived results from original paper!
Reproducing derived results from original paper!

But still simplified likelihood scans, not the full statistical model.
Given a probability model $p(X|\theta)$ and a data $x_0$

- The **likelihood function** is a function of the parameter $\theta$, and the value is given by $L(\theta) = p(X = x_0 | \theta)$

- But $L(\theta)$ doesn’t describe the distribution in $X$

- Technically the likelihood function doesn’t have enough enough information to generate synthetic data (toy Monte Carlo), which is needed for most frequentist statistical procedures

Colloquially, the term likelihood function is used in HEP often when we mean the full probability model $p(X|\theta)$

- We should be clear in paper, but no need to belabor the point here

**Note:** an intermediate provide a function $f(x, \theta) \propto p(x | \theta)$ (e.g. a NN likelihood ratio)

- it is a function of the data $x$, but you can’t (easily) sample from it (e.g. to generate toy MC)
Parameters modify rates only?

- yes
- no

Analysis is number counting?

- yes
- no

Grid for signal is available?

- yes
- no

Need continuous parametrization?

- yes
- no

- yes
- no

Interpolation of signal needed

- yes
- no

Replace original signal with new signal in following

- yes
- no

HEPData:
Tables of rates, acceptances, and systematic variations

HistFactory XML
for signal and background (stored in HEPData?)

RooFit/RooStats workspace

Shapes based on binned Templates?
Parameters modify rates only?

Analysis is number counting?

Grid for signal is available?

Need continuous parametrization?

RECAST and/or fast simulation to create grid points

Interpolation of signal needed

Replace original signal with new signal in following

HEPData: Tables of rates, acceptances, and systematic variations

Shapes based on binned Templates?

HistFactory XML for signal and background (stored in HEPData?)

RooFit/RooStats workspace
Likelihood Publishing + RECAST = TWO GREAT TASTES THAT TASTE GREAT TOGETHER.
Binned vs. unbinned

The RooWorkspace was designed to be able to store any type of statistical model → source of many complications

broadly we have two classes of analyses: binned and unbinned
Recent progress

Recently ATLAS has started publishing full likelihoods to HEPData for SUSY and exotics searches

- Perfect for STXS-type measurements
More published pyhf probability models

More on the way

Published Likelihoods

Updating list of HEPData entries for publications using HistFactory JSON likelihoods:

- Search for squarks and gluinos in final states with same-sign leptons and jets using 139 fb$^{-1}$ of data collected with the ATLAS detector. 2020. doi:10.17182/hepdata.91214.
- Search for direct production of electroweakinos in final states with one lepton, missing transverse momentum and a Higgs boson decaying into two b-jets in (pp) collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. 2020. doi:10.17182/hepdata.90607.
- Search for bottom-squark pair production with the ATLAS detector in final states containing Higgs bosons, b-jets and missing transverse momentum. 2019. doi:10.17182/hepdata.89408.

Cite these!
From yesterday

Note the pros and cons

• This list of what is needed looks a lot like the discussions in reinterpretation of BSM searches ↓

WHAT IS NEEDED

• Information on experimental cut flows, efficiencies
• Information on backgrounds
• Information on results and corresponding correlations (becoming standard)
• Information on the likelihood
• Desirable to have results at particle level, and distributions (STXS or fiducial distr.)

Different analysis strategies

- Highly optimised analyses targeting specific properties / operators
  - “best possible” sensitivity
  - very model specific

- Fiducial and differential cross section measurements
  - minimise model dependence
  - relatively restricted sensitivity (hard to combine different channels)
  - re-interpretable outside experiment

- Differential measurements in experimentally sensitive observables per production mode (STXS)
  - model dependence from production mode definition
  - easy combination of different Higgs decay channels → sensitivity to large number of EFT operators
  - re-interpretable outside experiment

Makes me think of this ↓ from 2012
Gold Standard: Full Likelihoods

Different analysis strategies

- Highly optimised analyses targeting specific properties / operators
  → “best possible” sensitivity
  → very model specific

- Fiducial and differential cross section measurements
  → minimise model dependence
  → relatively restricted sensitivity (hard to combine different channels)
  → re-interpretable outside experiment

- Differential measurements in experimentally sensitive observables per production mode (STXS)
  → model dependence from production mode definition
  → easy combination of different Higgs decay channels → sensitivity to large number of EFT operators
  → re-interpretable outside experiment

For EFT measurements, I think combined fits based on the full likelihood models in the “folded” data space are the most principled approach

- Likelihood can be based on STXS or dedicated analysis
- This should be the gold standard for the flagship EFT results
The STXS combination measurement

Aim: EFT interpretation of the 139 fb^{-1} combination of H→ZZ*→4\ell, H→γγ and H→bb merged stage-1.2 STXS measurement

Mostly split into p_{T}^{H} categories (and n_{jet})

43 categories
26 categories
12 categories
5 categories

[see also talk by Davide Mungo]

Brian Moser

SMEFT Higgs Measurements with ATLAS
RECAST + STXS overcomes model dependence

The model dependence in STXS mainly connected to how results are conveyed.

- The phase space regions are just phase space regions, they don’t assume any model
- Paired with RECAST one could reinterpret any model using the STXS phase space regions

Different analysis strategies

- Highly optimised analyses targeting specific properties / operators
  - “best possible” sensitivity
  - very model specific

- Fiducial and differential cross section measurements
  - minimise model dependence
  - relatively restricted sensitivity (hard to combine different channels)
  - re-interpretable outside experiment

- Differential measurements in experimentally sensitive observables per production mode (STXS)
  - model dependence from production mode definition
  - easy combination of different Higgs decay channels → sensitivity to large number of EFT operators
  - re-interpretable outside experiment

Saskia Falke  Inputs to EFT fits  29/10/2020  3 / 22
Likelihoods don’t address reinterpretation for signals with different final states or kinematics (eg. Exotic Higgs)
\[ \mathcal{L}_{\text{SM}} = \frac{1}{4} W_{\mu \nu} W^{\mu \nu} - \frac{1}{4} B_{\mu \nu} B^{\mu \nu} - \frac{1}{4} G_{\mu \nu} G^{\mu \nu} \\
+ \bar{L} \gamma^{\mu} \left( i \partial_{\mu} - \frac{1}{2} g \tau^{a} W_{\mu}^{a} - \frac{1}{2} g' Y B_{\mu} \right) L + \bar{R} \gamma^{\mu} \left( i \partial_{\mu} - \frac{1}{2} g' Y B_{\mu} \right) R \\
+ \frac{1}{2} \left( \Phi - \frac{1}{2} g \tau^{a} W_{\mu}^{a} - \frac{1}{2} g' Y B_{\mu} \right) \Phi^{*} - V(\Phi) \\
\]

Kinetic energies and self-interactions of the gauge bosons

Kinetic energies and electroweak interactions of fermions

\[ \left| \left| \left( i \partial_{\mu} - \frac{1}{2} g \tau^{a} W_{\mu}^{a} - \frac{1}{2} g' Y B_{\mu} \right) q \right| \right|_{2} - V(q) \]

Interactions between quarks and gluons

Boson masses and couplings to Higgs

Model B rejected?
Recently we’ve made enormous progress in preserving the full analysis chain.

- Makes it possible to run a new new signal through the full analysis chain
ATLAS has started using RECAST to reinterpret SUSY and exotics searches

- Also relevant for exotic BSM Higgs scenarios
Shifting from reproducibility to reuse

• Reuse provides a forward-looking narrative, while reproducibility often perceived as backward-looking

• Reproducibility is a byproduct!

• Analysis Preservation distinct from reproducibility

• Helps with onboarding

• Empowers reuse, remixing, reproducibility

• Improves efficiency & equity

Open is not enough

Xiaodi Chen1, Sünje Dallmeier-Tiessen2, Robin Dasler3, Sebastian Feger3, Pawel Fokionos1, Jose Benito Gonzalez1, Harri Hirvonsalo1,4,5, Dinos Kousidis1, Artemis Lavasa1, Salvatore Mele1, Diego Rodriguez Rodriguez1, Tibor Simko1, Tim Smith1, Ana Trisovic1,6, Anna Tzicisnka1, Ioannis Tsanaktidis1, Markus Zimmermann2, Kyle Cramer6, Lukas Heinrich6, Gordon Watts6, Michael Hilgert6, Lara Lloret Iglesias6, Kati Lassila-Perini6 and Sebastian Neubert6*

The solutions adopted by the high-energy physics community to foster reproducible research are examples of best practices that could be embraced more widely. This first experience suggests that reproducibility requires going beyond openness.

Corrected: Publisher Correction

https://doi.org/10.1038/s41567-018-0342-2
Encouraging response by the community

Participants in Analysis Preservation Bootcamp showing off their ability to reproduce an LHC analysis. Photo Credit: Samuel Meehan

Instructors Danika MacDonnel and Giordon Stark working with participants. Photo Credit: Samuel Meehan.