Systematic Uncertainties in Particle Physics Analyses

Louis Lyons
Imperial College London and Oxford
l.lyons@physics.ox.ac.uk

FCC Workshop
Feb 2022

Systematics

Systematic Uncertainties relevant to almost every Particle Physics Analysis

Especially important at various stages of experiment:
(i) Funding application {e.g. FCC}
(ii) Early running of new detector/accelerator (e.g. FCC)
(iii) When large integrated luminosity collected (e.g. FCC)
{See Karl Jacob's talk on ECFA's WG1-PREC}

Usually involves much more work than estimating statistical uncertainties

Sometimes less objective than statistical uncertainties

Subject of PHYSTAT-Systematics remote meeting in November 2021

PHYSTAT = series of Workshops & Seminars on statistical issues in Particle Physics
PHYSTAT Homepage:
https://espace.cern.ch/phystat/_layouts/15/start.aspx#/SitePages/Home.aspx
(Sign up for notifications of future events)

PHYSTAT-Systematics

Little time for detailed discussion, so mainly advert to look at slides & videos of all talks

Topic: How systematics are incorporated in Particle Physics analyses

Reason: Estimating magnitude of systematics (e.g. integrate luminosity is ±2%)
is very specific for each systematic
Not so amenable to help from statisticians.

Workshop webpage: https://indico.cern.ch/event/1051224/
Slides and videos of all talks, plus discussions
Lots of background reading material:

- **Basic introduction via simple pendulum —> g**
- **Reviews of systematics**
- **Topics for discussion (2 versions)**

STATISTICIANS

Vital component of PHYSTAT Workshops
(Unfortunately miss informal discussions)

Largest number of Statisticians at PHYSTAT meeting
Statistician Speakers:

- David van Dyk: Statistics Introduction
- Brad Efron: Bootstrap
- Mikael Kuusela: Unfolding
- Short talks: Berger, Lockhart, Manole, Meng, Wasserman
- Responders: Brazzale, Davison, Schafer, Wasserman, (Junk)
- Sara Algeri: Statistician's summary

Structure of meeting

INTRODUCTORY TALKS:
Louis Lyons (Particles)
David van Dyk (Statistics)

PHYSICS AREAS
LHC1 - Event selection, Discovery, Limits: Lukas Heinrich
LHC2 - Precision measurements: Sasha Glazov
Dark Matter: Knut Mora
Neutrinos: Christophe Bronner
Flavour Physics: Tom Blake

ANALYSIS TOPICS:
Unfolding: Mikael Kuusela
Mis-Modelling: Adinda de Wit
Theory : Frank Tackmann
Bootstrap: Brad Efron
Machine Learning: Kyle Cranmer
"Publish your Likelihood": Kyle Cranmer

STATISTICIANS' SHORT TALKS
Jim Berger: Bayesian reflection on Systematics
Richard Lockhart: Systematic effects versus nuisance parameters
Tudor Manole: Background modelling: Optimal Transport v Classifier Extrapolation
Xiao-Li Meng: Partially Bayes: Inter-perspective approach re nuisance parameters
Larry Wasserman: Universal Inference

SUMMARY TALKS
Nick Wardle (Physicist)
Sara Algeri: (Statistician)

Structure of meeting

INTRODUCTORY TALKS:
Louis Lyons (Particles)
David van Dyk (Statistics)

PHYSICS AREAS
LHC1 - Event selection, Discovery, Limits: Lukas Heinrich **Most relevant for FCC**
LHC2 - Precision measurements: Sasha Glazov **Most relevant for FCC**
Dark Matter: Knut Mora
Neutrinos: Christophe Bronner
Flavour Physics: Tom Blake **Most relevant for FCC**

ANALYSIS TOPICS:
Unfolding: Mikael Kuusela
Mis-Modelling: Adinda de Wit
Theory : Frank Tackmann
Bootstrap: Brad Efron
Machine Learning: Kyle Cranmer
"Publish your Likelihood": Kyle Cranmer

STATISTICIANS' SHORT TALKS
Jim Berger: Bayesian reflection on Systematics
Richard Lockhart: Systematic effects versus nuisance parameters
Tudor Manole: Background modelling: Optimal Transport v Classifier Extrapolation
Xiao-Li Meng: Partially Bayes: Inter-perspective approach re nuisance parameters
Larry Wasserman: Universal Inference

SUMMARY TALKS
Nick Wardle (Physicist)
Sara Algeri: (Statistician)

Structure of meeting

INTRODUCTORY TALKS:

Louis Lyons (Particles): Good, Bad and Ugly

David van Dyk (Statistics): Difference between Pragmatic Approach and Full Likelihood **

PHYSICS AREAS

LHC1 - Event selection, Discovery, Limits: Lukas Heinrich. **

LHC2 - Precision measurements; Sasha Glazov

Dark Matter: Knut Mora: White paper, PCL, Maximum gap

Neutrinos: Christophe Bronner

Flavour Physics: Tom Blake P'_5 angular variable for $B^0 \rightarrow K^{*0}\mu\mu$

ANALYSIS TOPICS:

Unfolding: Mikael Kuusela: Narrow bins for unfolding, then combine

Mis-Modelling: Adinda de Wit: Template morphing; 'Discrete Profiling', Dauncey et al, arXiv:1408.6865

Theory : Frank Tackmann: 2 point systematics

Bootstrap: Brad Efron

Machine Learning: Kyle Cranmer: Many topics – future Workshop? **

"Publish your Likelihood": Kyle Cranmer: Recall first PHYSTAT in 2000

** = see following slides

STATISTICIANS' SHORT TALKS

Jim Berger: Bayesian reflection on Systematics

Richard Lockhart: Systematic effects versus nuisance parameters

Tudor Manole: Background modelling: Optimal Transport v Classifier Extrapolation

Xiao-Li Meng: Partially Bayes: Inter-perspective approach re nuisance parameters

Larry Wasserman: Universal Inference: Does not require regularity conditions.

SUMMARY TALKS

Nick Wardle (Physicist): Systematic review of Systematics

Sara Algeri: (Statistician): More active engagement of Statisticians in analyses

Simple example for Pragmatic versus Full Likelihood (How to include systematic in analysis)

Interpretation of David van Dyk's introductory talk at PHYSTAT-Systematics
https://indico.cern.ch/event/1051224/

Main experiment attempts to estimate physics param Φ
But also have to consider nuisance parameter ν
Subsidiary experiment also provides information on ν
e.g. Any Collider analysis, with ν as jet energy scale, electron ID, integrated L, etc.

Method 1: Pragmatic
Fix ν at best value → $\mathcal{L}_{main}(\Phi, \nu_{best})$
→ 'Best' value of param & Statistical uncert on Φ
Shift ν by σ_ν (as determined in subsid expt) → Systematic uncert on Φ

Method 2: Full
Use product of likelihoods from Main and Subsid measurements →
Statistical and Systematic uncertainties on Φ

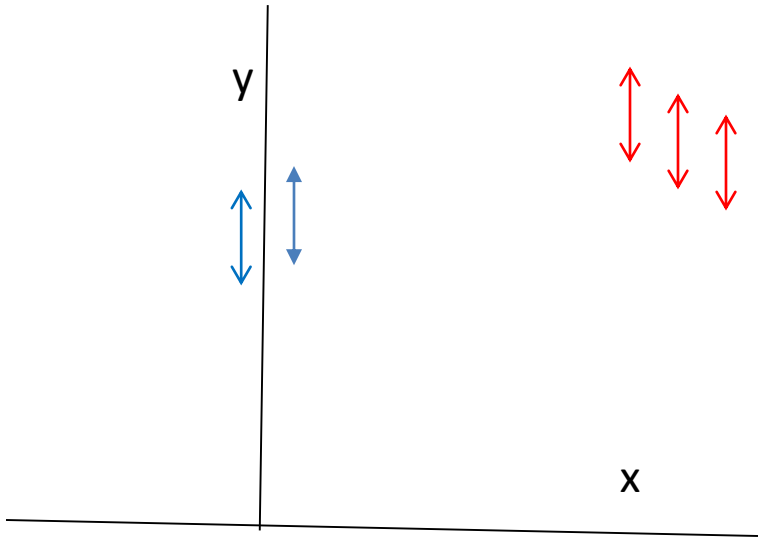
Use very simple example, for insight on difference. (Gaussian uncertainties)

Fit straight line **y = a + bx** to hits in detector planes

b = gradient = POI , a = intercept =nuisance param

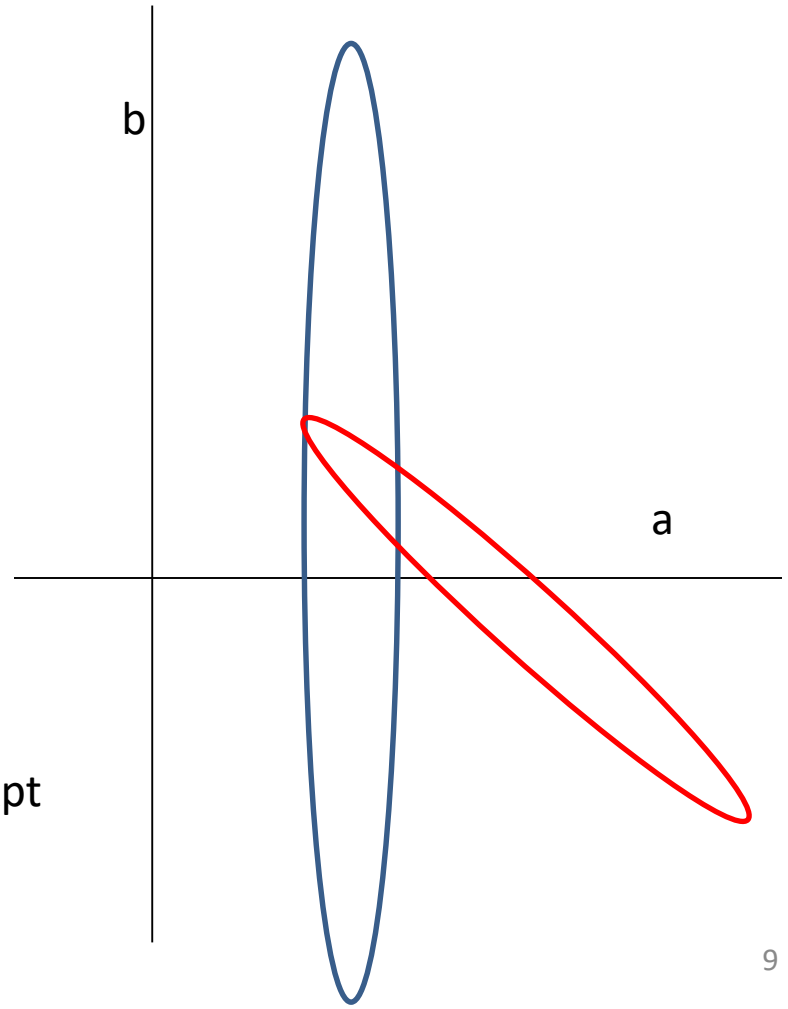
↕ = Main expt

↕ = Subsid expt



M = Inv Cov Matrix

$$M_{red} = \begin{pmatrix} A & C \\ C & B \end{pmatrix} \qquad M_{blue} = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}$$



Problems with 'Pragmatic'

1) 'Best' value of b unchanged by subsidiary expt
2) Overlarge systematic if subsid expt has large uncertainty

(Not much used in Particle Physics)

Lukas Heinrich:
Systematics at LHC for event selection, discovery and limits

Summary of building statistical models, including systematics

Interpolation for MC generated with different values of systematic parameters

Using ML to reduce dependence on systematics

'Ranking plots' for effect of systematics

Difference between expected sensitivity, before and after data taking

Need to check validity of asymptotic formulae

Publishing Probability Models → Giving Statisticians real data

Kyle Cranmer:
Four approaches to systematics with Machine Learning

"How many lectures on 'Systematics & ML?"
e.g. Assessing systematics when using ML
Systematics of ML procedure: Network structure, local minima, training samples
Using ML to reduce systematics.

Two notions of "incorporate":
• Don't be wrong: view analysis chain as fixed and propagate systematic uncertainty through it.
• Try to be "optimal": adjust the training of ML components so that the analysis is sensitive after accounting for systematics

Four approaches to Systematics
* Propagation of errors.
* Data augmentation
* Domain adaption
* Parameterized models

Propagating uncertainty with a pivotal classifier. Use adversarial network to pivot

Future

Still to be decided

Possibilities:

- Further Workshops e.g. At end of Run 3
  - Estimating Systematics
  - Systematics for ML
- Seminars e.g. Partons
  - Cosmology
  - Specific analyses

Your ideas welcome