Terascale Statistics School 2024

Report of Contributions

https://indico.desy.de/e/Statistics_2024

Contribution ID: **1** Type: **not specified**

Registration

Tuesday 2 April 2024 10:00 (1 hour)

Contribution ID: **2** Type: **not specified**

Welcome

Tuesday 2 April 2024 11:00 (5 minutes)

Presenters: MELZER-PELLMANN, Isabell (Deutsches Elektronen-Synchrotron DESY); BEHNKE, Olaf (CMS (CMS Fachgruppe TOP))

Contribution ID: **3** Type: **not specified**

Statistical Inference Lecture 1

Tuesday 2 April 2024 11:05 (1h 25m)

Plan for Lectures 1 and 2:
- Quick review of probability, frequentist vs. Bayesian approaches
- Parameter estimation, maximum likelihood, asymptotic properties of MLEs
- Hypothesis tests –general formalism, p-values
- Confidence intervals from inversion of test, from likelihood function
- General analysis including nuisance parameters, asymptotics

Presenter: COWAN, Glen (RHUL)

Contribution ID: **4** Type: **not specified**

Combine Tool Tutorial I

Tuesday 2 April 2024 14:00 (1h 30m)

Plan for Tutorial parts I - IV:
In this tutorial we will go through the main features of the Combine software package which provides a command-line interface to many different statistical techniques, available inside RooFit/RooStats, that are used widely inside CMS.
In the first part of the tutorial we'll explain the format of the Combine configuration file (datacard), and then discuss how to set up a simple counting experiment, followed by the shape analysis setup using ROOT histograms as inputs. Then we'll look at how to build simultaneous fits using independent categories, add systematic uncertainties and extract the limits or uncertainty estimates for the parameters of interest. Second part of the tutorial describes how to construct parametric models with RooFit and extract the measurements using Combine.

Presenters: NIGAMOVA, Aliya (University of Hamburg); CORMIER, Kyle (Uni Zurich); WARDLE, Nicholas

Contribution ID: **5** Type: **not specified**

Combine Tool Tutorial II

Tuesday 2 April 2024 16:00 (1h 30m)

Presenters: NIGAMOVA, Aliya (University of Hamburg); CORMIER, Kyle (Uni Zurich); WARDLE, Nicholas

Contribution ID: **6** Type: **not specified**

Statistics for Machine Learning Lecture I

Wednesday 3 April 2024 09:15 (1h 30m)

Plan for Lectures I - III:
LHC Physics with its vast amount of data and its precise first-principle simulations is a perfect case for using modern machine learning techniques. This includes data acquisition, analysis, simulations, and inference. I will introduce the main ML-concepts and ML-tools in a physics-specific manner, with a focus on statistics aspects. I will start by introducing Bayesian neural networks and likelihood loss functions and show how we can train classification networks with a probabilistic output. I will then introduce a range of generative networks, which are transforming not only our daily lives but also LHC physics. Finally, I will discuss how these ML-methods enable unfolding or inverse simulations and optimal inference for the LHC and more broadly. While there will be no tutorial for this lecture, all students are encouraged to deepen their understanding using the tutorials which are provided with the lecture notes https://arxiv.org/abs/2211.01421 Updated lecture notes are here:
https://www.thphys.uni-heidelberg.de/~plehn/pics/modern_ml.pdf

Presenter: PLEHN, Tilman (Heidelberg University)

Contribution ID: **7** Type: **not specified**

Statistical Inference Lecture II

Wednesday 3 April 2024 11:15 (1h 30m)

Presenter: COWAN, Glen (RHUL)

Contribution ID: **8** Type: **not specified**

Combine Tool Tutorial III

Wednesday 3 April 2024 14:15 (1h 15m)

Presenters: NIGAMOVA, Aliya (University of Hamburg); CORMIER, Kyle (Uni Zurich); WARDLE, Nicholas

Contribution ID: **9** Type: **not specified**

Combine Tool Tutorial IV

Wednesday 3 April 2024 16:00 (1h 30m)

Presenters: NIGAMOVA, Aliya (University of Hamburg); CORMIER, Kyle (Uni Zurich); WARDLE, Nicholas

Contribution ID: **10** Type: **not specified**

Statistics for Machine Learning II

Thursday 4 April 2024 09:00 (1h 30m)

Presenter: PLEHN, Tilman (Heidelberg University)

Contribution ID: **11** Type: **not specified**

Unfolding Lecture

Thursday 4 April 2024 11:15 (1h 30m)

In high-energy physics, unfolding is a critical statistical process for interpreting experimental data that is complicated by the intrinsic ill-posedness of the problem. This complexity arises from the need to provide heuristics for statistical estimates that disentangle true physical phenomena from observational distortions. We present a typical roadmap for why, when, and how unfolding is applied in high energy physics experiments and how the treatment of uncertainties influences considerations such as the choice of algorithm and regularisation.

Presenter: CROFT, Vincent

Contribution ID: **12** Type: **not specified**

Special Session on Uncertainties: Introductory Lecture and Lecture on Treatment of Asymmetric Uncertainties

Thursday 4 April 2024 14:00 (1h 30m)

What you mean by 'the error on the result', which is not as simple as you think.
- Bayesian and frequentist probability and confidence.
- The Neyman construction.
- Statistical and systematic uncertainties
- Asymmetric uncertainties and how to handle them

Presenter: BARLOW, Roger (Huddersfield)

Contribution ID: **13** Type: **not specified**

Special Session on Uncertainties: Lecture on Errors on Errors

Thursday 4 April 2024 16:00 (1h 30m)

A statistical model is presented where the uncertainty in assignment of systematic errors is taken into account. Estimates of the systematic variances are modeled as gamma distributed variables. The resulting confidence intervals show interesting and useful properties, e.g., when averaging measurements to estimate their mean, the size of the confidence interval increases as a for decreasing goodness-of-fit, and averages have reduced sensitivity to outliers. The basic properties of the model are presented and several types of examples relevant for Particle Physics are explored. (Based on G. Cowan, Eur. Phys. J. C (2019) 79:133, E. Canonero et al., Eur. Phys. J. C (2023) 83:1100.)

Presenter: COWAN, Glen (RHUL)

Contribution ID: **14** Type: **not specified**

Statistics of Machine Learning Lecture III

Friday 5 April 2024 09:15 (1h 30m)

Presenter: PLEHN, Tilman (Heidelberg University)

Contribution ID: **15** Type: **not specified**

School Summary

Friday 5 April 2024 11:15 (1 hour)

Looking back, a reminder of what you've learned in the week and, looking forward, how you're likely to use it.

Presenter: BARLOW, Roger (Huddersfield)

Contribution ID: **16** Type: **not specified**

Closing of School - Good Bye

Friday 5 April 2024 12:15 (15 minutes)

Presenters: MELZER-PELLMANN, Isabell (Deutsches Elektronen-Synchrotron DESY); BEHNKE, Olaf (CMS (CMS Fachgruppe TOP))