Contribution ID: **16** Type: **Oral**

Consistent multi-differential histogramming and summary statistics with YODA2

Thursday, 14 March 2024 15:30 (20 minutes)

In the contemporary landscape of advanced statistical analysis toolkits, ranging from Bayesian inference to machine learning, the seemingly straightforward concept of a histogram often goes unnoticed. However, the power and compactness of partially aggregated, multi-dimensional summary statistics with a fundamental connection to differential and integral calculus make them formidable statistical objects. Expressing these concepts robustly and efficiently in high-dimensional parameter spaces is a non-trivial challenge, especially when the resulting library is meant to remain usable by scientists rather than software engineers.

A decade after its initial release, the YODA statistical library has been redesigned from the ground, aiming to generalise its principles while addressing real-world usage requirements in the era of expanding computational power and vast datasets. We will summarise the core principles required for consistent generalised histogramming and outline some of the C++ metaprogramming techniques adopted to handle dimensionality relationships in the revamped YODA histogramming library. Used both in Rivet and Contur, YODA is a key component of physics data–model comparison and statistical interpretation in collider physics.

Significance

The YODA library is a key component in the Rivet and Contur packages. 10 years after its initial release, YODA has been redesigned using modern C++ techniques to provide generalised histogramming in arbitrary dimensions and addressing various other shortcomings of the initial release series.

Experiment context, if any

References

https://arxiv.org/abs/2312.15070

Primary authors: BUCKLEY, Andy (University of Glasgow (GB)); GUTSCHOW, Christian (UCL (UK)); YELLEN, Jamie (University of Glasgow); YEH, Yoran (University College London (UK))

Presenter: GUTSCHOW, Christian (UCL (UK))

Session Classification: Track 2: Data Analysis - Algorithms and Tools

Track Classification: Track 2: Data Analysis - Algorithms and Tools