Contribution ID: **76**　　　　　　　　　　　　　　　　　　　　　　　Type: **Plenary**

Open Questions in Statistical Practice for Particle Physics

Friday 23 August 2024 10:00 (30 minutes)

Statistical methods play a crucial role in the data analysis and interpretation in particle physics experiments. As we deepen in the theory and build more complex experiments, new challenges arise in statistical practice. In this talk, we explore some of the open questions, new and old, and ongoing debates in statistical methodology as applied to particle physics research.

Hypothesis testing is one of the basic grounds in particle physics research: when claiming a discovery, when confirming a given theoretical model or maybe simply to convince ourselves that our detector simulation properly reproduces data. Despite being a rather basic and classical concept in statistics, common practices are not always fully aligned with statistical theory. We are still asking ourselves why requiring for "5-sigma" to present an excess as a discovery or what does it really imply. We started to get used to jargon like global and local significance or "look-elsewhere effect", but what do they really mean?

The explosion of machine learning provides at the same time new powerful techniques, but also poses new problems in its interaction with classical statistics. We are familiar with its use in optimal classification, but more and more applications are appearing ranging from alternatives to Montecarlo simulation to detector effects unfolding, among others.

Optimal combination of results from different experiments is also a complex task when accounting for correlations and systematics. A crucial part of this problem lies in the amount of information that should be made publicly available to permit an optimal combination or a correct and efficient contrast with theory.

Furthermore, we discuss the challenges derived from properly incorporating systematic uncertainties into the statistical models. This is especially challenging for non-gaussian systematics such as the so-called asymmetric errors or the those derived from theory.

It becomes more and more frequent the use of unfolding to correct experimental results for detector effects and provide a direct way of comparison with theory, yet its application comes with inherent limitations and complexities. Challenges arise from the interplay of detector effects, statistical fluctuations, and model assumptions, leading to uncertainties in the unfolded results. Additionally, unfolding methods must contend with the non-trivial task of balancing between bias and variance, often requiring sophisticated algorithms and careful validation procedures.

Through an examination of these open questions and challenges, this talk aims to stimulate discussion and bring awareness of the assumptions and approximation lying behind the different methods.

Primary author:　MATORRAS, Francisco (Instituto de Fisica de Cantabria, Santander, IFCA (ES))

Presenter:　MATORRAS, Francisco (Instituto de Fisica de Cantabria, Santander, IFCA (ES))

Session Classification:　Plenary

Track Classification:　H: Statistical Methods for Physics Analysis in the XXIst Century