

INSS Statistics Project Summary

This mini-project explores properties of statistical tests used to discover a signal process by counting a number n of events and for each event measuring a quantity x that follows a given distribution under the hypothesis that the event is signal or background. The hypothesis that all of the events in the sample are of the background type is tested; if this is rejected at a sufficiently high level of significance then one can claim discovery of the signal.

The project begins with some simple “paper-and-pencil” exercises to investigate the properties of the statistical test of the hypothesis that all events are background. This is followed by a numerical calculation of the expected significance with which one can reject the background-only hypothesis if one were to simply count the events in a given range of the variable x (the search region). In the final (more challenging) part of the project, one needs to write a short Monte Carlo program to carry out the same tests but using not the number of events found in a given range of x but rather based on all of the x values from the entire event sample. Key concepts are the distinction between the discovery significance from a given data sample and the *expected* (or median) discovery significance under the hypothesis of a given signal model, the latter quantity is used, e.g., to optimize an analysis.

The materials for the project can be found at

<http://www.pp.rhul.ac.uk/~cowan/stat/mainz2018/project/>

Some useful background information can be found in Ref. [1]

References

- [1] Glen Cowan, *Statistics for Searches at the LHC*, in Einan Gardi, Nigel Glover and Aidan Robson (eds.) Proceedings of the 69th Scottish Universities Summer School in Physics, St. Andrews, August-September 2012, Scottish graduate series, Springer (2015); e-print arXiv:1307.2487 [hep-ex].