2.4 Distributions

2.4.1 Scenario: Working with histograms

Paul studies journal publications about a specific track *observable a*, in order to find data on its measured distribution, *dn/da*. He finds an article that publishes this distribution as measured by an experiment during the first year of running. The data are presented in a table, showing for each *a bin* the value of *dn/da* normalized to total number of measured events and the corresponding error. As the observable is predicted to have a Gaussian distribution, Paul tries to *fit* it with a Gaussian shape, but uncertainties are too high. After some more search, he finds a newer article by the same experiment that publishes the same distribution, but measured during the second and third year of data taking. Luckily, both distributions are presented with the same binning, and despite they are normalized to a different total number of events, Paul is able to add up the two distributions after simply deducing the original number of entries in each bin, and the total number of events. The combined distribution has smaller uncertainties and produces a better overall fit.

2.4.2 Literature

1. J. Orear, Notes on Statistics for Physicists (UCRL-8417, 1958)
2. D. J. Hudson, Lectures on Elementary Statistics and Probability (CERN 63-29, 1963)
3. W.T. Eadie et al, Statistical methods in experimental physics (North-Holland, 1971 and Elsevier, 1983)
4. S.L. Meyer, Data Analysis for Scientists and Engineers (Wiley, 1975)
5. S. Brandt, Statistical and computational methods in data analysis (North-Holland, 1976)
6. A. Frodesen et al, Probability and Statistics in Particle Physics (Oxford University Press, 1979)
7. L. Lyons, Statistics for Nuclear and Particle Physicists (Cambridge University Press 1986, reprinted 1989 and 1999)
8. R. Barlow, Statistics: A Guide to the Use of Statistical Methods in the Physical Sciences (Wiley, 1989)
9. B. Roe, Probability and Statistics in Experimental Physics (Springer, 1997)
10. J. Taylor, An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements (University Science Books, 1997)
11. G. Cowan, Statistical Data Analysis (Oxford Science Publications 1998)

2.4.3 Aims of the material

The goal of this scenario is to familiarize students with the concept of histograms, which is the key tool in physics analysis. Understanding of histograms and manipulations with them requires good knowledge of the basics of statistics, and in particular of binomial, Poisson and Gaussian distributions. Though it is not explicitly suggested by the scenario, the students are expected on this stage to identify the necessity of re-visiting the basics of e.g. Poisson distribution. In case of difficulties in doing this, the teacher should provide the necessary hints.

The students are not expected to study all the suggested reference literature, as practically every book offers a good description of the subject.

In the end, it could be considered an important achievement if every student will identify a favourite book or two on statistics, for further reference.