Empirical Bayes Unfolding of Elementary Particle Spectra at the Large Hadron Collider

Mikael Kuusela & Victor M. Panaretos

Section de Mathématiques
Ecole Polytechnique Fédérale de Lausanne

Technical Report # 01/14, January 2014
(Submitted for Publication, Under Revision)

Chair of Mathematical Statistics

Abstract

We consider the so-called unfolding problem in experimental high energy physics, where the goal is to estimate the true spectrum of elementary particles given observations distorted by measurement error due to the limited resolution of a particle detector. This an important statistical inverse problem arising in the analysis of data at the Large Hadron Collider at CERN. Mathematically, the problem is formalized as one of estimating the intensity function of an indirectly observed Poisson point process. Particle physicists are particularly keen on unfolding methods that feature a principled way of choosing the regularization strength and allow for the quantification of the uncertainty inherent in the solution. Though there are many approaches that have been considered by experimental physicists, it can be argued that few – if any – of these deal with these two key issues in a satisfactory manner. In this paper, we propose to attack the unfolding problem within the framework of empirical Bayes estimation: we consider Bayes estimators of the coefficients of a basis expansion of the unknown intensity, using a regularizing prior; and employ a Monte Carlo expectation-maximization algorithm to find the marginal maximum likelihood estimate of the hyperparameter controlling the strength of the regularization. Due to the data-driven choice of the hyperparameter, credible intervals derived using the empirical Bayes posterior lose their subjective Bayesian interpretation. Since the properties and meaning of such intervals are poorly understood, we explore instead the use of bootstrap resampling for constructing purely frequentist confidence bands for the true intensity. The performance of the proposed methodology is demonstrated using both simulations and real data from the Large Hadron Collider.

Keywords: *Poisson Inverse Problems; High Energy Physics; Uncertainty Quantification; Poisson Process; Regularization; Bootstrap; Monte Carlo EM Algorithm.*