Multidimensional integration and covariance matrix

C. Palmisano and G. Gervino

Dipartimento di Fisica, Universitá degli Studi di Torino, Italy

In this paper we study the possibility to integrate numerically a multivariate normal distribution, with given mean and covariance matrix, over 16-, 32-, 64-dimensional spaces, by mean of the ellipsoidal nested sampling method (a multidimensional integration technique). In other words, we consider the problems of identifying the most appropriate model for a given physical system and of assessing the model contribution to the measurement uncertainty. These problems are studied in terms of Bayesian model selection and model averaging. As the evaluation of the "evidence" Z (i.e. the integral of Likelihood x Prior over the space of the measurand and the parameters) becomes impracticable when this space has 20 - 30 dimensions or even more, then it is mandatory to elaborate an appropriate numerical strategy. Among the many algorithms for calculating Z, we focus on the ellipsoidal nested sampling, which is a technique based on three pillars: the study of the iso-likelihood contour lines of the integrand, a probabilistic estimate of the volume of the parameter space contained within the iso-likelihood contours and the random samplings from hyperellipsoids embedded in the integration variables. Using the points collected by the chosen algorithm, i.e. the list of the discharged likelihoods, we also obtained the post-data distributions and the relevant samples.

The consistency of the covariance matrices, which are calculated by the discharged likelihoods or by the samples using the input covariance matrices and the aspected Wishart distribution, has also been investigated.

As possible test of the obtained results, we applied the present statistical method to gamma-ray spectroscopy with HpGe detectors, that is a common technique in many fields such as nuclear physics, radiochemistry, nuclear medicine and neutron activation analysis. The use of HpGe detectors is chosen in situations where isotope identification is needed because of their excellent resolution. Our challenge is to obtain the "best" spectroscopy data possible in every measurement situation, i.e. to extract the most physical information. In the present case "best" is a combination of statistical (number of counts) and spectral quality (peak, width and position) over a wide range of counting rates. In this framework, we applied Bayesian methods and the Ellipsoidal Nested Sampling, in order to study the most likely distribution for the shape of HpGe spectra. In treating these experiments, the prior information suggests to model the likelihood function with a product of Poisson distributions. We present and discuss the efforts that have been done in order to optimize the statistical methods to HpGe detector outputs with the aim to evaluate to a better order of precision the detector efficiency, the absolute measured activity and the spectra background.