Results-sheet for random Seed = 10

March 21, 2024

Figure 1: Problem of N events with one jet per event, which can be either from class c or b. The problem is addressed with the four models: (from top to bottom) Gaussian Process, Dirichlet, Unimodal and Point-estimate. Left and center columns indicate the priors and posteriors on the c- and b-distributions, respectively. The right column shows the prior and posterior for the fractions of events with a c-jet. Although one can see a slight improvement in the Gaussian Process and Dirichlet models, none of the above solutions yields a satisfactory posterior. As we show in next Section, this outcome can be improved with many jets per event since the dependence between them at the event-by-event level provides profitable information for the inference process.

Figure 2: Problem of N events with four jets per event, which can be from either of the classes *cccc*, *ccbb* or *bbbb*. Figure shows the learning stages using a Dirichlet model. First line corresponds to the prior knowledge on the *c*- and *b*-distributions (left), and fraction (uniform) of each one of the classes in the sample (right). Red cross in the simplex indicates the true value for the classes fractions. The following lines correspond to the posterior knowledge of the same properties after having seen $N = 100,\ 250$ and 500 events, respectively. See text for the *'Actual data'* and for the binning in the simplex plots.

Figure 3: Learning flavor mixtures for Gaussian Process model. Same details as in Fig. 2.

Figure 4: Learning flavor mixtures for Unimodal model. Same details as in Fig. 2.

Figure 5: Learning flavor mixtures for Point Estimate model. Same details as in Fig. 2.

Figure 6: Summary plot for all models posterior convergence in c- and b-distributions in GN1. See text for details.

Figure 7: Summary plot for all models posterior convergence to the classes fraction in the samples. Observe that in contrast to Fig. 6, in this plot each point refers to only one parameter (the classes fraction) and not an average over many (each bin in the c- and b-distributions), therefore we have a more fluctuating plot. These fluctuations are still larger with small $N = 100$. In any case, it can be recognized a pattern in which the performance improves with data (N) and with including more prior-knowledge info in the models. Here the horizontal axis indicates the ratio of posterior to prior probabilities in the bin where lies the true fraction.