MDL exercises, seventh handout

(due Monday April 13, 14.00)

1. The prequential plug-in code based on the ML estimator is a universal code. Below we only consider one-dimensional models such that the data are i.i.d. according to all distributions in the model. For such models, there is usually a *minimal sample size* x^k such that the ML estimator is defined, and the log-loss obtained by predicting a new outcome based on the ML estimator of the past cannot be infinite, as soon as the first k outcomes x^k have been observed. In this exercise you may ignore the contribution of the initial sample x^k to any codelengths you compute. Just pretend that the contribution of these is zero. Thus, for a Poisson sequence we always have $k = 1$ and the regret of an i.i.d. Poisson sequence x^n becomes the prequential codelength of outcomes $x_2, \ldots, x_n$ minus the codelength for outcomes $x_2, \ldots, x_n$ based on the ML estimator for all outcomes.

   (a) Consider the prequential ML code for the Poisson model. Show that the regret can be infinite.

   (b) Now we consider possible extensions (of $n > k$ outcomes) of a sequence of one outcome, restricting ourselves to *i.i.d.* models (sets of distributions). Below, by 'the order of the extension', we mean 'the order of data points $x_{k+1}, \ldots, x_n$', i.e. any permutation of these points gives a different 'order'. Prove or give counterexamples to the following claims.

      i) There is no model for which the prequential ML codelength depends on the order of the extension.

      ii) The prequential ML codelength depends on the order of the extension for all 1-dimensional models.

      iii) There is no model for which the Bayesian marginal codelength depends on the order of the extension.

      iv) The Bayesian marginal codelength depends on the order of the extension for all models.

2. a) Show that for the Poisson model, when $n = 1$ (1 outcome), the minimax regret is infinite (Hint: use Stirling's approximation!) and hence the NML distribution is undefined. For bonus points, you may also show that the NML is also undefined for data of arbitrary fixed length (all $n > 1$).

   b) Explain why, for the Poisson model restricted to parameters $\mu \in [1, 100]$ and data of length $n = 1$, the minimax regret is finite and hence NML is defined. For bonus points, you may once again show that the same holds for any fixed $n > 1$. *(for 2c see backside!)*

c) Show that for the Poisson model restricted to parameters $\mu \in [1, 100]$ for all large enough n, there is a sequence of data for which the prequential plug-in code has regret that is *smaller* than the regret of the NML code by a term of order $(1/2) \log n$. Does this show that prequential plug-in code is really preferable over the NML code?