Solution Of Homework 2

January 30, 2011

1. **Exercise 9.1**

   Denote by n_i the number of elements in the sequence $y_1, y_2, \ldots, y_n$ that are equal to i. Clearly $\sum_i n_i = n$, and we can define the "empirica distribution" to be $g(i) = n_i/n$. The cumulative log loss of the expert f is

   $$-\sum_i n_i \log f(i) = -n \sum_i \frac{n_i}{n} \log f(i) = -n \sum_i g(i) \log f(i)$$

   To find the distribution f_i that minimizes this loss we use the Lagrange method for constrained minimization, which gives us

   $$-n \sum_i g(i) \log f(i) + \lambda \sum_i f(i)$$

   Taking the derivative with respect to $f(j)$ we get $g(j)/f(j) = \lambda$. For which the only solution is $f = g$, i.e. $f(i) = n_i/n$. The cumulative loss of this expert is

   $$-n \sum_i g(i) \log g(i) \doteq nH(g)$$

   Where $H(g)$ is the entropy of the empirical distribution g.

2. **Exercise 9.2**

   If we know the probability that each horse wins and the odds on each horse, we can compute the expected reward for better a dollar on the i'th horse, which is $p_i o_i$. To maximize our gain we should put all of our money on a horse with the maximal expected reward.

   In section 9.3 the sequence of horse wins is not assumed to be stochastic, let alone that we know the probabilities. Instead, we have access to the bet distributions of other gamblers, some of which hopefully know what they are doing. Our goal is not to do much worse than the gambler that performed best in hindsite.

3. **Exercise 9.3**

   As $n \geq \log_2 N$, $2^n \geq N$. Consider the 2^n binary sequences of length n. Suppose each expert assigns a probability of one to one of these sequences, and zero to all of the others.

   From Theorem 9.1 we know that

   $$V_n(\mathcal{F}) = \ln \sum_{x_n \in \mathcal{Y}^n} \sup_{f \in \mathcal{F}} f_n(x_n)$$

   In our construction there are exactly N sequences x_n for which there is a corresponding expert that assigns probability 1, the rest of the $2^n - N$ sequences are assigned probability zero by all of the experts. The result is that $V_n(\mathcal{F}) = \ln N$

4. **Exercise 9.5**

$$\begin{aligned}
&\sum_{y^n \in Y^n} q(y^n) \ln \frac{\sup_{f \in \mathcal{F}} f_n(y^n)}{\hat{P}_n(y^n)} - \sum_{y^n \in Y^n} q(y^n) \ln \frac{\sup_{f \in \mathcal{F}} f_n(y^n)}{q(y^n)} \\
=& \sum_{y^n \in Y^n} q(y^n) \ln \frac{q(y^n)}{\hat{P}_n(y^n)} = - \sum_{y^n \in Y^n} q(y^n) \ln \frac{\hat{P}_n(y^n)}{q(y^n)} \\
\geq& -\ln \sum_{y^n \in Y^n} q(y^n) \frac{\hat{P}_n(y^n)}{q(y^n)} = \ln \sum_{y^n \in Y^n} \hat{P}_n(y^n) = \ln 1 = 0
\end{aligned}$$

So

$$\sum_{y^n \in Y^n} q(y^n) \ln \frac{\sup_{f \in \mathcal{F}} f_n(y^n)}{\hat{P}_n(y^n)} \geq \sum_{y^n \in Y^n} q(y^n) \ln \frac{\sup_{f \in \mathcal{F}} f_n(y^n)}{q(y^n)}$$

$$* * ** \tag{1}$$