Algorithmic Statistics Problem Set 3

Problem 1 Consider a Gaussian version of the broadcast tree model. We have a complete d-ary tree of depth R and some parameter $0 < \alpha < 1$. The value at the root is sampled from $X_0 \sim N(0,1)$. Each of its children $X_1, \dots, X_d$ is then sampled independently according to the distribution

$$X_i \sim \alpha X_0 + \sqrt{1-\alpha^2} \cdot N(0,1)\,.$$

We then repeat this process on each of the children of $X_1, \dots, X_d$ and so on until we have sampled values at all nodes in the tree. As in the broadcast tree model shown in class, we will consider the problem of predicting the value at the root given only observations at the leaves (the nodes at depth R).

1. Given only the values at the leaves of the tree, say $Z_1, \dots, Z_{d^R}$, explicitly compute the posterior distribution of X_0 in terms of $Z_1, \dots, Z_{d^R}$

2. We say that nontrivial prediction for X_0 is possible if there exists a predictor f that is a function of the values at the leaves such that

$$\mathbb{E}[(X_0 - f(Z_1, \dots, Z_{d^R}))^2] < 1$$

   where the expectation is over all of the randomness in the tree. Determine the critical value α_0 for α in terms of d. Prove that as $R \to \infty$, when $\alpha > \alpha_0$ then nontrivial prediction is always possible and when $\alpha < \alpha_0$ then nontrivial prediction becomes impossible.

Problem 2 Suppose you have sample access to a distribution D supported on $[N]$, and you can interact with a prover who knows D. For this problem we will go back to uniformity testing: the goal is to use as few samples as possible from D, as well as a polynomial amount of computation with the prover, to decide whether D is the uniform distribution or $TV(D, \text{uniform}) \geq 0.01$. (This is the same model Tal Herman discussed in his guest lecture on Oct 15 – see lecture recording for clarification.)

Prove that in any such prover-verifier protocol where (a) if D is uniform the verifier says "yes" with probability at least $2/3$ and (b) if D is far from uniform, no matter how the prover behaves the verifier says "no" with probability at least $2/3$, the verifier must draw $\Omega(\sqrt{N})$ samples from D. Hint: you may appeal without proof to the uniformity testing lower bound from lecture 1.