Homework 5

10-704 Information Processing and Learning

Instructor: Aarti Singh

The HW is worth 40 (+5 bonus) pts and is **due on May 9 at noon**. Hand in to: Michelle Martin GHC 8001. If she is not around, note down the time on your HW sheet and slide it under her door.

1. [10 pts] *Large Deviations:* A tissue sample injected with a fluoroscent dye emits photons according to an iid Poisson process with rate of λ photons per second. What is the probability that in n seconds a microscope has observed an average of μ photons per second, assuming no scattering or other loss? (Hint: Even though a Poisson process is not finite alphabet, you can use Sanov's theorem since the polynomial term in Sanov's upper bound can be dropped if the set E is convex. Also, your answer should be in terms of λ, μ and n only.)

2. [10 pts] *Error exponent for universal codes:* Recall that a universal code of rate R can describe every iid source with entropy $H \leq R$ with probability of error asymptotically tending to 0. In this problem, we will use Sanov's theorem to calculate the probability of error exponent if a universal code of rate R is used to encode a source with distribution Q.

   (a) Note that for a non-universal source code, an error occurs if the sequence is not typical, i.e. empricial entropy deviates from true entropy by more than $\epsilon \to 0$. Based on this, when does an error occur if a universal code of rate R is used? What is the set E we would use for Sanov's theorem?

   (b) Assume $\epsilon = 0$. Find P^* in terms of Q and R.

   (c) Comment on the cases when $H(Q) > R$ and $H(Q) < R$.

3. [10 pts] *Bayes optimal test:* Show that the Bayesian probability of error is minimized by the aposteriori ratio test:

$$\frac{P_0(x^n)Pr(H_0)}{P_1(x^n)Pr(H_1)} \underset{H_1}{\overset{H_0}{\gtrless}} 1$$

4. [10 pts] *GLRT:* Compute the GLRT test statistic for the following hypothesis testing problem, where σ^2 and X ($n \times d$ matrix) are known and β ($d \times 1$ vector) is unknown:

$$H_0 : Y \sim \mathcal{N}(0, \sigma^2 I_{d\times d})$$
$$H_1 : Y \sim \mathcal{N}(X\beta, \sigma^2 I_{d\times d})$$

   Argue that the GLRT test statistic is essentially an energy detector in the projected subspace spanned by the columns of X. You may assume $n \geq d$. What happens when $n < d$?

   [+5 pts] Bonus Question: Can you propose a hypothesis test that might work if $n < d$?