Homework 5

36-705

Due: Thursday October 1st by 3pm.

1. Let $\mathcal{A}$ consist of all sets of the form

$$A = [a_1, b_1] \bigcup [a_2, b_2] \bigcup [a_3, b_3]$$

   where $a_1 \leq b_1 \leq a_2 \leq b_2 \leq a_3 \leq b_3$. Find the VC dimension of $\mathcal{A}$.

2. Suppose that $X_1, \ldots, X_n \sim N(\mu, \mu^2)$. Compute the likelihood function. Find a minimal sufficient statistic.

3. Suppose that we have $X_1, \ldots, X_n \sim N(\mu, \sigma^2)$. Let

$$\widehat{\theta} = X_1,$$

   and consider the sufficient statistic $T(X_1, \ldots, X_n) = \sum_{i=1}^n X_i$.

   (a) Compute the Rao-Blackwellized estimator.
   Hint: Use the fact that $(X_1, \sum_{i=1}^n X_i)$ has a joint Gaussian distribution. Compute its mean and covariance.

   (b) Compute the risk of $\widehat{\theta}$ and compare it to the risk of the Rao-Blackwellized estimator.

   (c) Suppose you instead tried to use the estimator $\mathbb{E}[X_1|X_2]$. What is wrong with this estimator?

4. Let $X_1, \ldots, X_n \sim U[\theta, 1+\theta]$. Show that

$$T(X_1, \ldots, X_n) = (\min_i X_i, \max_i X_i)$$

   is a minimal sufficient statistic.

5. For the following distributions, check that they are in an exponential family by appropriately re-writing their density. List the sufficient statistics and canonical parameters.

   (a) The Gamma distribution with shape parameter k and scale θ.

   (b) The central χ^2 distribution with k degrees of freedom.

   (c) The multinomial distribution (treat the number of trials n as fixed).

   In each case, also compute the log-partition function A and differentiate it to obtain the expected value of the sufficient statistics.