ECES 632 Homework 2

1. What is the definition of a positive semi-definite matrix? Are the diagonal elements of a positive semi-definite matrix guaranteed to be ≥ 0, if so, why? Use this to explain why an efficient estimator (i.e. one which obtains the Cramér-Rao lower bound) must be UMVU.

2. Write the CRLB for the variance of an unbiased an estimator for $\boldsymbol{\theta}$ operating on N independently and identically distributed observations $\mathbf{R}_{\mathtt{i}}$ from $\mathsf{p}_{\mathbf{R};\boldsymbol{\theta}}(\mathbf{r};\boldsymbol{\theta})$ in terms of the CRLB for the variance of an unbiased estimator of $\boldsymbol{\theta}$ operating on 1 observation $\mathbf{R}_0$.

3. An electronic thermometer monitors the temperature τ at an unspecified location. The voltage reading V at time k from the thermometer is well modeled as the sum of τ with a zero-mean Gaussian distributed noise with known variance σ^2, so that

$$V_{\mathtt{k}} := \tau + n_{\mathtt{k}}$$

Suppose that we independently measure and record the voltage reading from the thermometer for times $\mathtt{k} \in \{1, \ldots, \mathtt{K}\}$, so that we have the vector $\mathbf{v} := [v_1, v_2, \ldots, v_{\mathtt{K}}]$ and we would like to estimate the temperature τ.

   (a) Write the statistical model for this estimation problem, identifying the observations and parameters.

   (b) Prove that $\sum_{\mathtt{k}=1}^{\mathtt{K}} V_{\mathtt{k}}$ is a sufficient statistic.

   (c) Is $\hat{\tau}(\mathbf{V}) := \frac{1}{\mathtt{K}} \sum_{\mathtt{k}} V_{\mathtt{k}}$ an unbiased estimator for the temperature τ?

   (d) Write the CRLB for the variance of an unbiased estimator of τ from $\mathbf{v}$.

   (e) Is $\hat{\tau}(\mathbf{V}) := \frac{1}{\mathtt{K}} \sum_{\mathtt{k}} V_{\mathtt{k}}$ UMVU? Is it an efficient estimator for the temperature?

4. Download the neural spike train data from the Signal Processing Information Base (SPIB) at Rice University `http://spib.rice.edu/spib/data/signals/bio/spiketrain.html`.[1] The elements of this vector of data are described as : *Intervals between discharges recorded from a single neuron responding to an acoustic stimulus. Measurements consist of the time intervals (measured in seconds) between successive action potentials. Such data are frequently modeled as point processes.* Load the data into MATLAB by navigating to the directory into which you extracted the spike train data, and typing `load spiketrain.mat`. Let's start by taking a look at a histogram of the data by typing `hist(spiketrain,50);`. The shape appears roughly to match the shape of an exponential distribution. Let us explore how well such a distribution can fit this data. Hypothesize the following model for the data

$$X_{\mathtt{k}} \sim \text{ i.i.d. } \mathsf{Expo}(\lambda)$$

where the random variable $X_{\mathtt{k}}$ takes on the observed value `spiketrain(k)`. Our goal is to estimate λ, the parameter for the exponential distribution.

   (a) Identify the observations, the parameter, and the statistical model for this problem from the discussion above.

   (b) Is there an efficient estimator for λ (prove your answer)? (Hint, use the continuation of the CRLB.)

   (c) Write a closed form expression for the maximum likelihood estimator for λ by setting the derivative of the log likelihood function equal to zero and solving for λ.

   (d) Use the expression derived in the previous problem to write a MATLAB script which calculates the maximum likelihood estimate for λ from the spike train data. Turn in both the MATLAB code and the calculated estimate with your homework.

[1] You will probably want to download the matlab format data, which is gzip compressed. Email me if you don't have a gzip decompressor and I'll send you the decompressed MAT file.