EE621 ADAPTATION AND LEARNING **Instructor: Ali H. Sayed**
Homework Assignment #2 (Inference and Regression) **Due: March 25, 2019**

---

1) A random variable $\boldsymbol{y}$ is uniformly distributed over the interval $0 \leq y \leq a$. We observe N independent and identically distributed realizations $\{\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_N\}$ and wish to determine the maximum-likelihood estimate for a.
   (a) Verify that
   $$f(y_1, \ldots, y_N; a) = \frac{1}{a^N}\,\mathbb{I}\left[0 \leq \max_{1\leq n\leq N} y_n \leq a\right]$$
   where $\mathbb{I}[x]$ denotes the indicator function. Conclude that $\widehat{a}_{\text{ML}} = \max_{1\leq n\leq N} y_n$.
   (b) Show that the estimator is biased, namely, establish that $\mathbb{E}\,\widehat{\boldsymbol{a}}_{\text{ML}} = \left(\frac{N}{N+1}\right)a$.
   (c) Show that $T(y) = \max_{1\leq n\leq N} y_n$ is a sufficient statistic for a.
2) Consider two coins, one with probability p for "heads" and the other with probability q for "heads". The first coin is chosen with probability α and the second coin is chosen with probability $1-\alpha$. When a coin is chosen, it is tossed and the outcome is noted as either $y_n = 1$ (for heads) or $y_n = 0$ (for tails) in the $n-$th experiment. Given a collection of N measurements $\{y_n,\, n = 1, 2, \ldots, N\}$, write down the EM recursions for estimating the parameters $\{\alpha, p, q\}$.
3) A person is choosing colored balls, one at a time and with replacement, from one of two urns. Each urn has red (R) and blue (B) balls. The proportion of red balls in urn #1 is 70%, while the proportion of blue balls in urn #2 is 40%. Either urn can be selected with equal probability. Four balls are selected and the following colors are observed: RBBR. It is not known which urn generated each of the balls.
   (a) What is the likelihood of observing this outcome?
   (b) What would be the most likely color for the ball at $n = 5$?
   (c) Can you estimate the sequence of urn selections for these four time instants?
4) The state of some system of interest is described by a binary variable $\boldsymbol{\theta}$, which can be either 0 or 1 with equal probability. Let $\boldsymbol{y}$ be a random variable that is observed according to the following probability distribution:

	$\boldsymbol{y}=0$	$\boldsymbol{y}=1$
$\boldsymbol{\theta}=0$	q	$1-q$
$\boldsymbol{\theta}=1$	$1-q$	q

   We collect N independent observations $\{\boldsymbol{y}(1), \ldots, \boldsymbol{y}(N)\}$. We wish to employ these observations in order to learn about the state of the system. Assume the true state is $\boldsymbol{\theta} = 0$.
   (a) Find the optimal mean-square-error estimator for $\boldsymbol{\theta}$, namely, $\widehat{\boldsymbol{\theta}}_N = \mathbb{E}\,(\boldsymbol{\theta}|\boldsymbol{y}(1), \boldsymbol{y}(2), \ldots, \boldsymbol{y}(N))$.
   (b) Assume $q \neq 0.5$. Show that $\widehat{\boldsymbol{\theta}}_N$ decays to zero exponentially in the mean-square-error sense as $N \to \infty$. What happens when $q = 0.5$?
   (c) Find an expression for the variance of $\widehat{\boldsymbol{\theta}}_N$. Find its limit as $N \to \infty$. Why are these results useful?
5) Consider N independent and identically-distributed real-valued random variables, $\{\boldsymbol{y}(n),\ n = 0, 1, \ldots, N-1\}$. Each $\boldsymbol{y}(n)$ has a Gaussian distribution with zero mean and variance σ^2. We use the observations $\{\boldsymbol{y}(n)\}$ to estimate σ^2 using $\widehat{\boldsymbol{\sigma}}^2 = \alpha \sum_{n=0}^{N-1} \boldsymbol{y}^2(n)$, for some scalar parameter α to be determined.
   (a) What is the mean of the estimator $\widehat{\boldsymbol{\sigma}}^2$ in terms of α and σ^2?
   (b) Evaluate the mean-square-error below in terms of α and σ^2, namely, m.s.e. $= \mathbb{E}\,(\widehat{\boldsymbol{\sigma}}^2 - \sigma^2)^2$.
   (c) Determine the optimal scalar α that minimizes the m.s.e.. Is the corresponding estimator biased or unbiased?
   (d) For what value of α would the estimator be unbiased? What is the m.s.e. of this estimator and how does it compare to the m.s.e. of the estimator from part (c)?
6) Let $w^\star$ and $w^\star_{\text{reg}}$ denote the solutions to the following problems:
   $$w^\star \;\stackrel{\Delta}{=}\; \underset{w\in\mathbb{R}^M}{\text{argmin}}\; \|d - Hw\|^2, \qquad w^\star_{\text{reg}} \;\stackrel{\Delta}{=}\; \underset{w\in\mathbb{R}^M}{\text{argmin}}\; \left\{\rho\|w\|^2 + \|d - Hw\|^2\right\}$$
   where $\rho > 0$. Let $Q = H^{\mathsf{T}}H$ and assume Q is invertible. Show that $w^\star_{\text{reg}} = (I + \rho Q)^{-1} w^\star$.

Computer project. The statement of the computer project is left purposefully vague to give you freedom to explore and choose the settings for your simulations. You are required to submit plots and your code, along with a description of your solution and commentary on the results.

Simulate the EM procedure applied to a Gaussian mixture model of at least 4 components. Include in your code a step to estimate the number of Gaussian components in the data. Explain this step and provide references if applicable.