Homework 4

Due: Saturday, Oct-24-2020, 5pm – Gradescope entry code: 948XVG

Please upload your answers timely to Gradescope. Start a new page for every problem. For the programming/simulation questions you can use any reasonable programming language. Comment your source code and include the code and a brief overall explanation with your answers.

1. In example 3 in class, we computed the MAP detector by reducing the problem to example 2. Compute the MAP detector by directly computing the likelihood ratio.

2. In class we defined a sufficient statistic \( V = g(Y) \) for detecting \( X \) as satisfying \( X \) and \( Y \) are independent conditional on \( V \). Another definition for \( g(Y) \) to be a sufficient statistic is that the likelihood ratio \( \Lambda(y) \) can be written as a function of \( g(y) \), i.e. there exists a function \( h \) such that \( \Lambda(y) = h(g(y)) \).
   
   (a) Explain why the second definition implies that one can perform optimal detection of \( X \) using only \( V \) without keeping \( Y \).

   (b) For example 2 covered in class, show that \( V = h^Ty \) satisfies both definitions of a sufficient statistic.

   (c) For Exercise 8.9(b), find a sufficient statistic for the detection based on \( (Y_1, Y_2) \), and check that it satisfies both definitions of a sufficient statistic.

3. Exercise 8.10 in text.

4. Exercise 8.11 part (a) in text.

5. In this problem we are going to implement the MAP rule for example 3 taught in class. Download the dataset from the course website [http://web.stanford.edu/class/ee278/homeworks/hw4-data.zip](http://web.stanford.edu/class/ee278/homeworks/hw4-data.zip). In example 3, we detect based on an observation \( Y \in \mathbb{R}^n \) following the statistical model

\[
Y = hX + gW + Z,
\]

where \( g \) and \( h \) are normalized to have norm 1, \( n = 10, W \sim \mathcal{N}(0, \sigma^2_w), Z \sim \mathcal{N}(0, \sigma^2_z I) \) and all vectors have the matching dimensions. We will further make the assumption that \( X \) is equally likely to be 0 or 1.

In practice, we may not know the parameters. So we need to first estimate the parameters and then estimate \( X \) from \( Y \).
a) To estimate the noise level $\sigma_z$, we first collect data in the form of $k = 3000$ observations which only contain noise $Z$ but no signal $hX$ nor interference $gW$. Propose an estimator for $\sigma_z^2$. The data are stored as column vectors in $Y_z$ in data.mat. Implement your method and report the estimated value of $\sigma_z^2$. You can assume the noise is independent across the 3000 observations.

b) We now want to further estimate $g$ and $\sigma_w$. We collect data in the form of another $k = 3000$ observations with both the interference $gW$ and the noise $Z$, but no signal $hX$. Propose an estimator for $\sigma_w^2$ and $g$. The data are stored as column vectors in $Y_w$ in data.mat. Implement your method and report the estimated value of $\sigma_w^2$ and $g$. You can assume that while the interference and the noise are independent across the 3000 observations, the parameters to estimate remain the same.

c) Finally, we need to estimate $h$. Propose an estimator for $h$ based on another $k = 3000$ observations which follows model (1). The observations are stored as column vectors in $Y_x$ in data.mat. Implement your method and report the estimated value of $h$. You can assume the signals are independent across the observations.

d) Based on the observations in part (c) and your estimated channel parameters, perform detection using the MAP rule on each of the 3000 observations. The true signals are stored in $X$ as an $1 \times n$ row vector in $X_label.mat$. Report the overall error rate. Your detector should not involve inverting a 10 by 10 matrix.

Please include your code. You can use any programming languages (e.g. Matlab, Python).
Exercise 8.10  (a) Consider a binary hypothesis testing problem, and denote the hypotheses as \( X = 1 \) and \( X = -1 \). Let \( \mathbf{a} = (a_1, a_2, \ldots, a_n) \) be an arbitrary real \( n \)-vector and let the observation be a sample value \( y \) of the rv \( Y = Xa + Z \), where \( Z \sim \mathcal{N}(0, \sigma^2 I_n) \) and \( I_n \) is the \( n \times n \) identity matrix. Assume that \( Z \) and \( X \) are independent. Find the ML decision rule and find the probabilities of error \( \Pr(e \mid X=0) \) and \( \Pr(e \mid X=1) \) in terms of the function \( Q(x) \).

(b) Now suppose a third hypothesis, \( X = 0 \), is added to the situation of (a). Again the observation rv is \( Y = Xa + Z \), but here \( X \) can take on values \(-1, 0, \) or \(+1\). Find a one-dimensional sufficient statistic for this problem (i.e., a one-dimensional function of \( y \) from which the likelihood ratios

\[
\Lambda_1(y) = \frac{p_{Y \mid X=1}(y \mid 1)}{p_{Y \mid X=0}(y \mid 0)} \quad \text{and} \quad \Lambda_{-1}(y) = \frac{p_{Y \mid X=-1}(y \mid -1)}{p_{Y \mid X=0}(y \mid 0)}
\]

can be calculated).

(c) Find the ML decision rule for the situation in (b) and find the probabilities of error, \( \Pr(e \mid X=x) \) for \( x = -1, 0, +1 \).

(d) Now suppose that \( Z_1, \ldots, Z_n \) in (a) are IID and each is uniformly distributed over the interval \(-2 \) to \(+2\). Also assume that \( \mathbf{a} = (1, 1, \ldots, 1) \). Find the ML decision rule for this situation.

Exercise 8.11  A sales executive hears that one of his salespeople is routing half of his incoming sales to a competitor. In particular, arriving sales are known to be Poisson at rate one per hour. According to the report (which we view as hypothesis \( X = 1 \)), each second arrival is routed to the competition; thus under hypothesis 1 the interarrival density for successful sales is \( f(y \mid X=1) = ye^{-y}; y \geq 0 \). The alternative hypothesis \( (X = 0) \) is that the rumor is false and the interarrival density for successful sales is \( f(y \mid X=0) = e^{-y}; y \geq 0 \). Assume that, a priori, the hypotheses are equally likely. The executive, a recent student of stochastic processes, explores various alternatives for choosing between the hypotheses; he can only observe the times of successful sales, however.

(a) Starting with a successful sale at time 0, let \( S_1 \) be the arrival time of the
ith subsequent successful sale. The executive observes \( S_1, S_2, \ldots, S_n \) \( (n \geq 1) \) and chooses the maximum a posteriori probability hypothesis given this observation. Find the joint probability density \( f(S_1, S_2, \ldots, S_n \mid X=1) \) and \( f(S_1, \ldots, S_n \mid X=0) \) and give the decision rule.