ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE
School of Computer and Communication Sciences

Principles of Digital Communications: Assignment date: March 7, 2012
Summer Semester 2012 Due date: March 14, 2012

Homework 4

Project: Form groups of 2 up to 4 people for the project. Send the names of the group members to saeid.haghighatshoar@epfl.ch by latest Wednesday 14th March.

Reading Part for the next week: Error Probability (Section 2.6) and Summary (Section 2.7). Read also the Definitions, Theorems, and Lemmas of the appendix Facts About Matrices (Appendix 2.A).

Problem 1. *(Fisher-Neyman Factorization Theorem)*

Consider the hypothesis testing problem where the hypothesis is $H \in \{0, 1, \ldots, m-1\}$, the observable is Y, and $T(Y)$ is a function of the observable. Let $f_{Y|H}(y|i)$ be given for all $i \in \{0, 1, \ldots, m-1\}$. Suppose that there are functions $g_1, g_2, \ldots, g_{m-1}$ so that for each $i \in \{0, 1, \ldots, m-1\}$ one can write

$$f_{Y|H}(y|i) = g_i(T(y))h(y). \tag{1}$$

1. Show that when the above conditions are satisfied, a MAP decision depends on the observable Y only through $T(Y)$. In other words, Y itself is not necessary.*(Hint: Work directly with the definition of a MAP decision rule.)*

2. Show that $T(Y)$ is a sufficient statistic, that is $H \to T(Y) \to Y$. *(Hint: Start by observing the following fact: Given a random variable Y with probability density function $f_Y(y)$ and given an arbitrary event $\mathcal{B}$, we have*

$$f_{Y|Y\in\mathcal{B}} = \frac{f_Y(y)1_{\mathcal{B}}(y)}{\int_{\mathcal{B}} f_Y(y)dy}. \tag{2}$$

Proceed by defining $\mathcal{B}$ to be the event $\mathcal{B} = \{y : T(y) = t\}$ and make use of (2) applied to $f_{Y|H}(y|i)$ to prove that $f_{Y|H,T(Y)}(y|i,t)$ is independent of i.)

For the following two examples, verify that condition (1) above is satisfied. You can then immediately conclude from part (1) and (2) above that $T(Y)$ is a sufficient statistic.

1. *(Example 1)* Let $Y = (Y_1, Y_2, \ldots, Y_n)$, $Y_k \in \{0, 1\}$, be an independent and identically distributed (i.i.d) sequence of coin tosses such that $P_{Y_k|H}(1|i) = p_i$. Show that the function $T(y_1, y_2, \ldots, y_n) = \sum_{k=1}^{n} y_k$ fulfills the condition expressed in equation (1).*(Notice that $T(y_1, y_2, \ldots, y_n)$ is the number of 1's in $y_1, y_2, \ldots, y_n$.)*

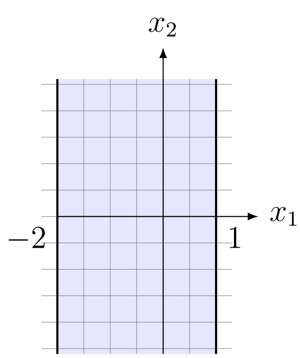
2. *(Example 2)* Under hypothesis $H = i$, let the observable Y_k be Gaussian distributed with mean m_i and variance 1; that is

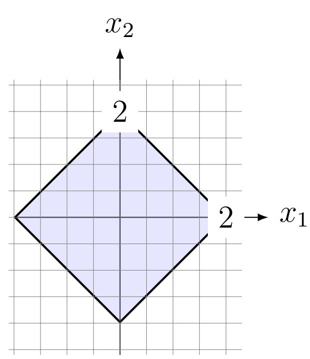
$$f_{Y_k|H}(y|i) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(y-m_i)^2}{2}},$$

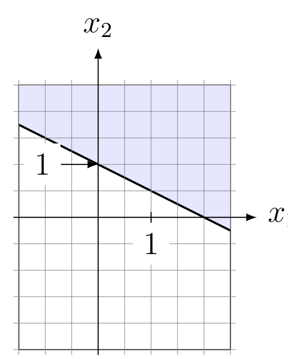
and $Y_1, Y_2, \ldots, Y_n$ be independently drawn according to this distribution. Show that the sample mean $T(y_1, y_2, \ldots, y_n) = \frac{1}{n} \sum_{k=1}^{n} y_k$ fulfills the condition expressed in equation (1).

Problem 2. *(Q-Function on Regions) [Wozencraft and Jacobs]*

Let $\boldsymbol{X} \sim \mathcal{N}(0, \sigma^2 I_2)$. For each of the three figures below, express the probability that $\boldsymbol{X}$ lies in the shaded region. You may use the Q-function when appropriate.

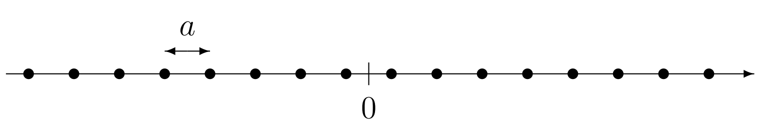


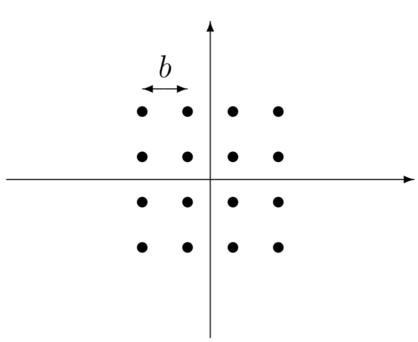




Problem 3. *(16-PAM versus 16-QAM)*

The following two signal constellations are used to communicate across an additive white Gaussian noise channel. Let the noise variance be σ^2.





Each point represents a signal $\mathbf{s_i}$ for some i. Assume each signal is used with the same probability.

1. For each signal constellation, compute the average probability of error, P_e, as a function of the parameters a and b, respectively.
2. For each signal constellation, compute the average energy per symbol, E_s, as a function of the parameters a and b, respectively:

$$E_s = \sum_{i=1}^{16} P_H(i) \parallel \mathbf{s_i} \parallel^{\mathbf{2}} \tag{3}$$

3. Plot P_e versus E_s for both signal constellations and comment.

Problem 4. *(Antenna Array)*

The following problem relates to the design of multi-antenna systems. The situation that we have in mind is one where one of two signals is transmitted over a Gaussian channel and is received through two different antennas. We shall assume that the noises at the two terminals are independent but not necessarily of equal variance. You are asked to design a receiver for this situation, and to assess its performance. This situation is made more precise as follows:

Consider the binary equiprobable hypothesis testing problem:

$$\begin{aligned} H=0 \ &: \ Y_1 = A + Z_1, \quad Y_2 = A + Z_2 \\ H=1 \ &: \ Y_1 = -A + Z_1, \quad Y_2 = -A + Z_2, \end{aligned}$$

where Z_1, Z_2 are independent Gaussian random variables with *different* variances $\sigma_1^2 \neq \sigma_2^2$, that is, $Z_1 \sim \mathcal{N}(0, \sigma_1^2)$ and $Z_2 \sim \mathcal{N}(0, \sigma_2^2)$. $A > 0$ is a constant.

1. Show that the decision rule that minimizes the probability of error (based on the observable Y_1 and Y_2) can be stated as

$$\sigma_2^2 y_1 + \sigma_1^2 y_2 \ \underset{1}{\overset{0}{\gtrless}} \ 0.$$

2. Draw the decision regions in the (Y_1, Y_2) plane for the special case where $\sigma_1 = 2\sigma_2$.
3. Evaluate the probability of the error for the optimal detector as a function of σ_1^2, σ_2^2 and A.

Problem 5. *(Sufficient Statistics)*

This exercise intends to elaborate more on the intuitive meaning of "sufficient statistic". As stated in the notes, let H be the hidden hypothesis that we want to discover. We are given the observable Y and our objective is to use Y to predict the value of H. However, sometimes the observable Y has some redundant information that is not really needed to predict H. More precisely, there is a function $T(Y)$ such that with $T(Y)$ we can do the same kind of prediction as with Y.

1. Let $H \in \{0, 1\}$ with $P_H(0) = P_H(1) = \frac{1}{2}$. Also, define the random variable $Y \in \{0, 1\}$ with the following law:

$$f_{Y|H}(y|0) = \begin{cases} \frac{1}{4}, & y = 0 \\ \frac{3}{4}, & y = 1. \end{cases}$$

   and

$$f_{Y|H}(y|1) = \begin{cases} \frac{3}{4}, & y = 0 \\ \frac{1}{4}, & y = 1. \end{cases}$$

   Now, suppose that we observe n i.i.d. samples $(Y_1, ..., Y_n)$ of Y. Show that $T(y_1, ..., y_n) = y_1 + ... + y_n$ is a sufficient statistic for H. **Hint:** try to write the quantity $f_{(Y_1,...,Y_n)|H}(\cdot|\cdot)$ in terms of $T(y_1, ..., y_n)$.

2. Let $Y_1, ..., Y_n$ be i.i.d. samples from the geometric distribution with parameter $H \in (0, 1)$. In other words, $P(Y_i = k) = (1 - H)^k H$ for $k = 0, 1, \ldots$. Show that $T(y_1, \ldots, y_n) = y_1 + \cdots + y_n$ is a sufficient statistic for predicting H.

3. Suppose that we have a binary hypothesis testing in which $H \in \{0, 1\}$ and $(Y_1, Y_2, \ldots, Y_n)$ is the observed random vector. Also assume that under hypothesis H the probability distribution of the observation is $f(y_1, y_2, \ldots, y_n|H)$, $H \in \{0, 1\}$.

   (a) Show that we can write $f(y_1, y_2, \ldots, y_n|H)$ as follows:

$$f(y_1, y_2, \ldots, y_n|H) = f(y_1, y_2, \ldots, y_n|0)L(y_1, y_2, \ldots, y_n)^H,$$

   where $L(y_1, y_2, \ldots, y_n) = \frac{f(y_1,y_2,\ldots,y_n|1)}{f(y_1,y_2,\ldots,y_n|0)}$ is the likelihood ratio.

   (b) Use Fischer-Neyman factorization theorem to show that $L(y_1, y_2, \ldots, y_n)$ is a sufficient statistics.

   (c) Using the result above, argue that why the MAP decision rule just depends on the likelihood ratio and why other details of the probability distributions are not important.