ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

School of Computer and Communication Sciences

Handout 8 Principles of Digital Communications

Problem Set 4 Mar. 17, 2021

PROBLEM 1. Consider a binary hypothesis testing problem specified by:

$$H=0\colon \begin{cases} Y_1 = Z_1 \\ Y_2 = Z_1 Z_2 \end{cases} \qquad H=1\colon \begin{cases} Y_1 = -Z_1 \\ Y_2 = -Z_1 Z_2, \end{cases}$$

where Z_1, Z_2, and H are independent random variables. Is Y_1 a sufficient statistic?

PROBLEM 2. We have seen that if $H \to T(Y) \to Y$, then the probability of error P_e of a MAP decoder that decides on the value of H upon observing both $T(Y)$ and Y is the same as that of a MAP decoder that observes only $T(Y)$. It is natural to wonder if the contrary is also true, specifically if the knowledge that Y does not help reduce the error probability that we can achieve with $T(Y)$ implies $H \to T(Y) \to Y$. Here is a counter-example. Let the hypothesis H be either 0 or 1 with equal probability. (The distribution of H is critical in this example.) Let the observable Y take four values with conditional probabilities

$$P_{Y|H}(y|0) = \begin{cases} 0.4 & \text{if } y=0 \\ 0.3 & \text{if } y=1 \\ 0.2 & \text{if } y=2 \\ 0.1 & \text{if } y=3 \end{cases} \qquad P_{Y|H}(y|1) = \begin{cases} 0.1 & \text{if } y=0 \\ 0.2 & \text{if } y=1 \\ 0.3 & \text{if } y=2 \\ 0.4 & \text{if } y=3 \end{cases}$$

and $T(Y)$ is the function

$$T(y) = \begin{cases} 0 & \text{if } y = \{0,1\} \\ 1 & \text{if } y = \{2,3\} \end{cases}$$

(a) Show that the MAP decoder $\hat{H}(T(y))$ that decides based on $T(y)$ is equivalent to the MAP decoder $\hat{H}(y)$ that operates based on y.

(b) Compute the probabilities $Pr\{Y=0|T(Y)=0, H=0\}$ and $Pr\{Y=0|T(Y)=0, H=1\}$. Is it true that $H \to T(Y) \to Y$?

PROBLEM 3. *(Fisher–Neyman factorization theorem)* Consider the hypothesis testing problem where the hypothesis is $H \in \{0,1,\dots,m-1\}$, the observable is Y, and $T(Y)$ is a function of the observable. Let $f_{Y|H}(y|i)$ be given for all $i \in \{0,1,\dots,m-1\}$. Suppose that there are positive functions $g_0, g_1, \dots, g_{m-1}, h$ such that for each $i \in \{0,1,\dots,m-1\}$ one can write

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

(a) 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.

(b) 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

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

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.

(c) (Example 1) Under hypothesis $H = i$, let $Y = (Y_1, Y_2, \ldots, Y_n) \in \{0,1\}^n$, be an independent and identically distributed 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 1s in y.

(d) (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 (1).

PROBLEM 4. Let X be a random variable with probability density function

$$p_X(x) = h(x)\exp(c(\theta)T(x) - B(\theta))$$

The set of probability distributions of this form is called the one-parameter *exponential family* distribution. For the rest of the problem assume $X_1, X_2, \ldots, X_n$ are i.i.d $\sim p_X$.

(a) Find a sufficient statistic for $(X_1, X_2, \ldots, X_n)$.
Hint: Find the joint pdf (or pmf) and use the Fisher-Neyman factorization theorem.

(b) Many distributions you know belong to this family e.g., Bernoulli, Poisson, Exponential, Laplace, Normal and Binomial distributions (with known number of trials).

Verify that the following distributions are members of the exponential family and, using your results from part (a), find a sufficient statistic for $(X_1, X_2, \ldots, X_n)$.

- Exponential Distribution: $p_X(x) = \lambda \exp(-\lambda x)\, \mathbb{1}\{x \geq 0\}$
- Laplace Distribution with known mean μ: $p_X(x) = \dfrac{1}{2\sigma}\exp\left(-\dfrac{|x-\mu|}{\sigma}\right)$
- Poisson Distribution: $p_X(x) = \dfrac{\lambda^x \exp(-\lambda)}{x!}$
- Binomial Distribution with known number of trials n: $p_X(x) = \binom{n}{x} p^x (1-p)^{n-x}$

PROBLEM 5. Consider a discrete memoryless channel (DMC). This is a channel model described by an input alphabet $\mathcal{X}$, an output alphabet $\mathcal{Y}$, and a transition probability[1] $P_{Y|X}(y|x)$. When we use this channel to transmit an n-tuple $x \in \mathcal{X}^n$, the transition probability is

$$P_{Y|X}(y|x) = \prod_{i=1}^{n} P_{Y|X}(y_i|x_i)$$

So far, we have come across two DMCs, namely the BSC (binary symmetric channel) and the BEC (binary erasure channel). The purpose of this problem is to see that for DMCs, the Bhattacharyya bound takes a simple form, in particular when the channel input alphabet $\mathcal{X}$ contains only two letters.

[1] Here we are assuming that the output alphabet is discrete. Otherwise we use densities instead of probabilities.

(a) Consider a transmitter that sends $c_0 \in \mathcal{X}^n$ and $c_1 \in \mathcal{X}^n$ with equal probability. Justify the following chain of (in)equalities:

$$\begin{aligned}
P_e &\overset{(a)}{\leq} \sum_y \sqrt{P_{Y|X}(y|c_0) P_{Y|X}(y|c_1)} \\
&\overset{(b)}{=} \sum_y \sqrt{\prod_{i=1}^n P_{Y|X}(y_i|c_{0,i}) P_{Y|X}(y_i|c_{1,i})} \\
&\overset{(c)}{=} \sum_{y_1,\ldots,y_n} \prod_{i=1}^n \sqrt{P_{Y|X}(y_i|c_{0,i}) P_{Y|X}(y_i|c_{1,i})} \\
&\overset{(d)}{=} \sum_{y_1} \sqrt{P_{Y|X}(y_1|c_{0,1}) P_{Y|X}(y_1|c_{1,1})} \\
&\quad \cdots \sum_{y_n} \sqrt{P_{Y|X}(y_n|c_{0,n}) P_{Y|X}(y_n|c_{1,n})} \\
&\overset{(e)}{=} \prod_{i=1}^n \sum_y \sqrt{P_{Y|X}(y|c_{0,i}) P_{Y|X}(y|c_{1,i})} \\
&\overset{(f)}{=} \prod_{a\in\mathcal{X}, b\in\mathcal{X}, a\neq b} \left(\sum_y \sqrt{P_{Y|X}(y|a) P_{Y|X}(y|b)} \right)^{n(a,b)},
\end{aligned}$$

where $n(a, b)$ is the number of positions i in which $c_{0,i} = a$ and $c_{1,i} = b$.

(b) The Hamming distance $d_H(c_0, c_1)$ is defined as the number of positions in which c_0 and c_1 differ. Show that for a binary input channel, i.e. when $\mathcal{X} = \{a, b\}$, the *Bhattacharyya bound* becomes

$$P_e \leq z^{d_H(c_0,c_1)},$$

where

$$z = \sum_y \sqrt{P_{Y|X}(y|a) P_{Y|X}(y|b)}$$

Notice that z depends only on the channel, whereas its exponent depends only on c_0 and c_1.

(c) Evaluate the channel parameter z for the following.

(i) The binary input Gaussian channel described by the densities

$$f_{Y|X}(y|0) = \mathcal{N}\left(-\sqrt{E}, \sigma^2\right)$$
$$f_{Y|X}(y|1) = \mathcal{N}\left(\sqrt{E}, \sigma^2\right)$$

(ii) The binary symmetric channel (BSC) with $\mathcal{X} = \mathcal{Y} = \{\pm 1\}$ and transition probabilities described by

$$P_{Y|X}(y|x) = \begin{cases} 1 - \delta, & \text{if } y = x \\ \delta, & \text{otherwise} \end{cases}$$

(iii) The binary erasure channel (BEC) with $\mathcal{X} = \{\pm 1\}$, $\mathcal{Y} = \{-1, E, 1\}$, and transition probabilities given by

$$P_{Y|X}(y|x) = \begin{cases} 1-\delta, & \text{if } y = x, \\ \delta, & \text{if } y = E, \\ 0, & \text{otherwise.} \end{cases}$$

Problem 6. Consider a QAM receiver that outputs a special symbol δ (called *erasure*) whenever the observation falls in the shaded area of the figure below, and does minimum-distance decoding otherwise. (This is neither a MAP nor an ML receiver.) Assume that $c_0 \in \mathbb{R}^2$ is transmitted and that $Y = c_0 + N$ is received where $N \sim \mathcal{N}(0, \sigma^2 I_2)$. Let P_{0i}, $i = \{0, 1, 2, 3\}$ be the probability that the receiver outputs $\hat{H} = i$ and let $P_{0\delta}$ be the probability that it outputs δ. Determine $P_{00}, P_{01}, P_{02}, P_{03}$, and $P_{0\delta}$.

Comment: If we choose $b - a$ large enough, we can make sure that the probability of error is very small (we say that an error occurred if $\hat{H} = i$, $i \in \{0, 1, 2, 3\}$ and $H \neq \hat{H}$). When $\hat{H} = \delta$, the receiver can ask for a retransmission of H. This requires a feedback channel from the receiver to the transmitter. In most practical applications, such a feedback channel is available.