ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

School of Computer and Communication Sciences

**Handout 8** Principles of Digital Communications

Problem Set 4 Mar. 11, 2015

Problem 1. *(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 $g_0, g_1, \dots, g_{m-1}$, and $h$ so 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:* Given a random variable $Y$ with probability density function $f_Y(y)$ and given an arbitrary event $\mathcal{B}$, we have

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

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 $i$.

Problem 2. *(Application of Factorization Theorem)* Using the result you proved in Problem 1, show the following:

(a) Under hypothesis $H = i$, let $Y = (Y_1, \dots, Y_n)$, $Y_i \in \{0, 1, 2, \dots\}$, be and i.i.d. sequence of Poisson random variables with parameter $\lambda_i > 0$. That is,

$$P_{Y_k|H}(y_k|i) = \frac{\lambda_i^{y_k}}{(y_k)!}e^{-\lambda_i}, \quad y_k \in \{0, 1, 2, \dots\}.$$

Show that $T(y_1, \dots, y_n) = \frac{1}{n}\sum_{i=1}^n y_i$ is a sufficient statistic. This statistic is called the *sample mean*.

(b) Under hypothesis $H_i$, the observable is $Y = (Y_1, \dots, Y_n)$ where $Y_k = \theta_i + Z_k$ and $Z_k$, $k = 1, 2, \dots, n$ are i.i.d. Exponential random variables with rate $\lambda > 0$, i.e.,

$$f_{Z_k}(z_k) = \begin{cases} \lambda e^{-\lambda z_k}, & \text{if } z_k \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

Show that the two-dimensional vector $T(y_1, \dots, y_n) = \left(\min_{k=1,\dots,n} y_k, \frac{1}{n}\sum_{k=1}^n y_k\right)$ is a sufficient statistic.

Problem 3. *(Sufficient Statistic)* Consider a binary hypothesis testing problem specified by:

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

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

*Hint:* If $Y = aZ$ for some scalar $a$ then $f_Y(y) = \frac{1}{|a|} f_Z(\frac{y}{a})$.

Problem 4. *(More on Sufficient Statistic)* We have seen that if $H \to T(Y) \to Y$ then the $P_e$ of a MAP decoder that observes both $T(Y)$ and $Y$ is the same as that of a MAP decoder that observes only $T(Y)$. You may wonder if the contrary is also true, namely if the knowledge that $Y$ does not help reducing the error probability that one can achieve with $T(Y)$ implies $H \to T(Y) \to Y$. Here is a counterexample. Let the hypothesis $H$ be either $0$ or $1$ with equal probability (the choice of distribution on $H$ is critical in this example). Let the observable $Y$ take four values with the following 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 following function

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

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

(b) Compute the probabilities $Pr(Y=0 \mid T(Y)=0, H=0)$ and $Pr(Y=0 \mid T(Y)=0, H=1)$. Do we have $H \to T(Y) \to Y$?

Problem 5. *(Repeat Codes and Bhattacharyya Bound)* Consider two equally likely hypotheses. Under hypothesis $H=0$, the transmitter sends $c_0 = (1,\ldots,1)^\mathsf{T}$ and under $H=1$ it sends $c_1 = (-1,\ldots,-1)^\mathsf{T}$, both of length $n$. The channel model is AWGN with variance $\sigma^2$ in each component. Recall that the probability of error for a ML receiver that observes the channel output $Y \in \mathbb{R}^n$ is

$$P_e = Q\left(\frac{\sqrt{n}}{\sigma}\right).$$

Suppose now that the decoder has access *only* to the sign of $Y_i$, $1 \le i \le n$, i.e., it observes

$$W = (W_1,\ldots,W_n) = (\mathrm{sign}(Y_1),\ldots,\mathrm{sign}(Y_n)). \tag{3}$$

(a) Determine the MAP decision rule based on the observable $W$. Give a simple sufficient statistic.

(b) Find the expression for the probability of error $\tilde{P}_e$ of the MAP decoder that observes $W$. You may assume that $n$ is odd.

(c) Your answer to (b) contains a sum that cannot be expressed in closed form. Express the Bhattacharyya bound on $\tilde{P}_e$.

(d) For $n = 1, 3, 5, 7$, find the numerical values of $P_e$, $\tilde{P}_e$, and the Bhattacharyya bound on $\tilde{P}_e$.

Problem 6. *(Bhattacharyya Bound for DMCs)* 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.

(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.

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

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

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) Compute $z$ for:

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

$$\begin{aligned} f_{Y|X}(y|0) &= \mathcal{N}(-\sqrt{E}, \sigma^2) \\ f_{Y|X}(y|1) &= \mathcal{N}(\sqrt{E}, \sigma^2). \end{aligned}$$

(ii) The Binary Symmetric Channel (BSC) with the 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 the 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 7. *(Tighter Union Bhattacharyya Bound: Binary Case)* In this problem we derive a tighter version of the *Union Bhattacharyya Bound* for binary hypotheses. Let

$$H = 0 : Y \sim f_{Y|H}(y|0) \qquad H = 1 : Y \sim f_{Y|H}(y|1).$$

(a) Argue that the probability of error of the MAP decision rule is

$$P_e = \int_y \min\left\{P_H(0)f_{Y|H}(y|0), P_H(1)f_{Y|H}(y|1)\right\} dy.$$

(b) Prove that for $a, b \geq 0$, $\min(a,b) \leq \sqrt{ab} \leq \frac{a+b}{2}$. Use this to prove the tighter version of *Bhattacharyya Bound*, i.e,

$$P_e \leq \frac{1}{2}\int_y \sqrt{f_{Y|H}(y|0)f_{Y|H}(y|1)}dy.$$

(c) Compare the above bound to that of Equation (2.18) in your course notes when $P_H(0) = \frac{1}{2}$. How do you explain the improvement by a factor $\frac{1}{2}$?