Handout 4 Principles of Digital Communications
Problem Set 2 Feb. 28, 2024

PROBLEM 1. Consider a binary hypothesis testing problem in which the hypotheses $H = 0$ and $H = 1$ occur with probability $P_H(0)$ and $P_H(1) = 1 - P_H(0)$, respectively. The observable Y takes values in $\{0, 1\}^{2k}$, where k is a fixed positive integer. When $H = 0$, each component of Y is 0 or 1 with probability $\frac{1}{2}$ and components are independent. When $H = 1$, Y is chosen uniformly at random from the set of all sequences of length $2k$ that have an equal number of ones and zeros. There are $\binom{2k}{k}$ such sequences.

(a) What is $P_{Y|H}(y|0)$? What is $P_{Y|H}(y|1)$?

(b) Find a maximum-likelihood decision rule for H based on y. What is the single number you need to know about y to implement this decision rule?

(c) Find a decision rule that minimizes the error probability.

(d) Are there values of $P_H(0)$ such that the decision rule that minimizes the error probability always chooses the same hypothesis regardless of y? If yes, what are these values, and what is the decision?

PROBLEM 2. Let us assume that a "weather frog" bases his forecast of tomorrow's weather entirely on today's air pressure. Determining a weather forecast is a hypothesis testing problem. For simplicity, let us assume that the weather frog only needs to tell us if the forecast for tomorrow's weather is "sunshine" or "rain". Hence we are dealing with binary hypothesis testing. Let $H = 0$ mean "sunshine" and $H = 1$ mean "rain". We will assume that both values of H are equally likely, i.e., $P_H(0) = P_H(1) = \frac{1}{2}$. For the sake of this exercise, suppose that on a day that precedes sunshine, the pressure may be modeled as a random variable Y with the following probability density function:

$$f_{Y|H}(y|0) = \begin{cases} A - \frac{A}{2}y, & 0 \le y \le 1 \\ 0, & \text{otherwise} \end{cases}$$

Similarly, the pressure on a day that precedes a rainy day is distributed according to

$$f_{Y|H}(y|1) = \begin{cases} B + \frac{B}{3}y, & 0 \le y \le 1 \\ 0, & \text{otherwise} \end{cases}$$

The weather frog's purpose in life is to guess the value of H after measuring Y.

(a) Determine A and B.

(b) Find the *a posteriori probability* $P_{H|Y}(0|y)$. Also find $P_{H|Y}(1|y)$.

(c) Show that the implementation of the decision rule $\hat{H}(y) = \arg\max_i P_{H|Y}(i|y)$ reduces to

$$\hat{H}_\theta(y) = \begin{cases} 0, & \text{if } y \le \theta \\ 1, & \text{otherwise} \end{cases}$$

for some threshold θ and specify the threshold's value.

(d) Now assume that the weather forecaster does not know about hypothesis testing and arbitrarily chooses the decision rule $\hat{H}_\gamma(y)$ for some arbitrary $\gamma \in \mathbb{R}$. Determine, as a function of γ, the probability that the decision rule decides $\hat{H} = 1$ given that $H = 0$. This probability is denoted $\Pr\left\{\hat{H}(Y) = 1 | H = 0\right\}$.

(e) For the same decision rule, determine the probability of error $P_e(\gamma)$ as a function of γ. Evaluate your expression at $\gamma = \theta$.

(f) Using calculus, find the γ that minimizes $P_e(\gamma)$ and compare your result to θ.

PROBLEM 3. [1] Let X be a discrete random variable such that

$$P\{X = n\} = \frac{2}{3^n} \forall n \in \mathbb{N}\backslash\{0\}.$$

We define the random variable Y as follows: knowing $X = n$, Y takes values n or $n+1$ with equal probability.

(a) Compute $E[X]$.

(b) Compute $E[Y|X = n]$ and deduce $E[Y|X]$, then $E[Y]$.

(c) Compute the joint probability of X and Y.

(d) Compute the marginal probability of Y.

(e) Compute $E[X|Y = i] (\forall i \in \mathbb{N}\backslash\{0\})$ and deduce $E[X|Y]$.

(f) Compute the covariance of X and Y.

PROBLEM 4. Consider testing two equally likely hypotheses $H = 0$ and $H = 1$. The observable $Y = (Y_1, \ldots, Y_k)^\mathsf{T}$ is a k-dimensional binary vector. Under $H = 0$ the components of the vector Y are independent uniform random variables (also called Bernoulli $\left(\frac{1}{2}\right)$ random variables). Under $H = 1$, the component Y_1 is also uniform, but the components Y_i, $2 \le i \le k$, are distributed as follows (i.e., $Y_1, \ldots, Y_k$ is a first-order Markov chain):

$$P_{Y_i|Y_1,\ldots,Y_{i-1}}(y_i|y_1, \ldots, y_{i-1}) = \begin{cases} 3/4, & \text{if } y_i = y_{i-1} \\ 1/4 & \text{otherwise} \end{cases}$$

(a) Find the decision rule that minimizes the probability of error.
Hint: Write down a short sample sequence $(y_1, \ldots, y_k)$ and determine its probability under each hypothesis. Then generalize.

(b) Give a simple sufficient statistic for this decision. (For the purpose of this question, a sufficient statistic is a function of y with the property that a decoder that observes y can not achieve a smaller error probability than a MAP decoder that observes this function of y.)

(c) Suppose that the observed sequence alternates between 0 and 1 except for one string of ones of length s, i.e. the observed sequence y looks something like

$$y = 01010101111111 \ldots 111111010101$$

What is the least s such that we decide for hypothesis $H = 1$?

[1] Probability Review Question from EPFL Applied Statistical Process 2017.

Problem 5. Consider the following binary hypothesis testing problem. The hypotheses are equally likely and the observable $Y = (Y_1, \ldots, Y_n)^\mathsf{T}$ is a n-dimensional real vector whose components are:

$$H_0 : Y_k = Z_k \qquad \text{versus} \qquad H_1 : Y_k = 2A + Z_k, \quad k = 1, \ldots, n,$$

where $A > 0$ is a positive constant and $Z_1, \ldots, Z_n$ is an i.i.d. noise sequence. In each of following cases, show that the MAP decision rule reduces to

$$\hat{H}(y_1, \ldots, y_n) = \begin{cases} 0 & \text{if } \sum_{k=1}^{n} \phi(y_i - A) < 0, \\ 1 & \text{otherwise,} \end{cases}$$

and find the function $\phi(\cdot)$.

(a) If Z_k are i.i.d. Gaussian noise samples with zero mean and variance σ^2, i.e.

$$f_{Z_k}(z_k) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{z_k^2}{2\sigma^2}}, \quad k = 1, \ldots, n.$$

(b) If Z_k are i.i.d. Laplacian noise samples with variance σ^2. That is,

$$f_{Z_k}(z_k) = \frac{1}{\sigma\sqrt{2}} e^{-\frac{\sqrt{2}}{\sigma}|z_k|}, \quad k = 1, \ldots, n.$$

Plot the noise densities for cases (a) and (b) for the same value of σ (take $\sigma = 1$ for convenience). Explain intuitively the difference of the functions $\phi(\cdot)$ that you found in (a) and (b).

Problem 6. Consider the binary hypothesis testing problem where the hypotheses are equally likely and the observable $Y = (Y_1, \ldots, Y_n)^\mathsf{T}$ is a n-dimensional real vector with components defined as

$$H_0 : Y_k = -A + Z_k \qquad \text{versus} \qquad H_1 : Y_k = A + Z_k, \quad k = 1, \ldots, n,$$

where $A > 0$ is a positive constant and $Z_1, \ldots, Z_n$ are i.i.d. Gaussian noise samples with variance σ^2. Find the decision rule that minimizes the probability of error. Compare your answer with that of Problem 5 part (a). What can you conclude?