ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

School of Computer and Communication Sciences

Principles of Digital Communications:	Assignment date: Feb 29, 2012
Summer Semester 2012	Due date: Mar 7, 2012

Homework 3

Reading for next week: From m-ary Hypothesis Testing (Section 2.2.2) up to and including Irrelevance and Sufficient Statistic (Section 2.5).

Problem 1. *(The "Wetterfrosch")*

Let us assume that a "weather frog" bases his forecast for 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. $P_H(0) = P_H(1) = \frac{1}{2}$.

Measurements over several years have led the weather frog to conclude that on a day that precedes sunshine the pressure may be modeled as a random variable Y with the following probability density function:

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

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

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

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

1. Determine A and B.

2. Find the *a posteriori probability* $P_{H|Y}(0|y)$. Also find $P_{H|Y}(1|y)$. *Hint:* Use Bayes' rule.

3. Plot $P_{H|Y}(0|y)$ and $P_{H|Y}(1|y)$ as a function of y.Show that the implementation of the decision rule $\hat{H}(y) = \arg\max_i P_{H|Y}(i|y)$ reduces to

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

for some threshold θ and specify the threshold's value. Do so by direct calculation rather than using the general result.

4. Now assume that you implement the decision rule $\hat{H}_\gamma(y)$ and determine, as a function of γ, the probability that the decision rule decides $\hat{H} = 1$ given that $H = 0$. This probability is denoted $Pr\{\hat{H}(y) = 1|H = 0\}$.

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

6. Using calculus, find the γ that minimizes $P_e(\gamma)$ and compare your result to θ. Could you have found the minimizing γ without any calculation?

Problem 2. *(Hypothesis Testing in Laplacian Noise)*

Consider the following hypothesis testing problem between two equally likely hypotheses. Under hypothesis $H = 0$, the observable Y is equal to $a + Z$ where Z is a random variable with Laplacian distribution

$$f_Z(z) = \frac{1}{2}e^{-|z|}.$$

Under hypothesis $H = 1$, the observable is given by $-a + Z$. You may assume that a is positive.

1. Find and draw the density $f_{Y|H}(y|0)$ of the observable under hypothesis $H = 0$, and the density $f_{Y|H}(y|1)$ of the observable under hypothesis $H = 1$.

2. Find the decision rule that minimizes the probability of error. Write out the expression for the likelihood ratio.

3. Compute the probability of error of the optimal decision rule.

Problem 3. *(Discrete Additive Gaussian Channel)*

The communication channel between a transmitter and a receiver is an AWG channel. If the input of the channel at time i is x then the output of the channel is $x + N_i$ where $N_i \sim \mathcal{N}(0, \sigma^2)$ is the additive noise of the channel. We assume that N_i is a sequence of i.i.d. random variables. We are going to send one bit of information through this channel in the following way : If the information bit is 1 we send A, n times through the channel. So the received sequence will be $Y_i = A + N_i$, $i = 1, \ldots, n$ whereas if the information bit is 0 we send $-A$, n times through the channel.

1. Write the problem as a hypothesis testing problem.

2. Derive the MAP rule assuming that 0 and 1 are equiprobable.

3. Find the error probability of the MAP rule and write it as a function of SNR $\triangleq \frac{A^2}{\sigma^2}$ and n. **Hint:** use the Q-Function defined as :

$$Q(x) \triangleq \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-\frac{t^2}{2}} dt.$$

4. Use the upper bound $Q(x) \leq \frac{1}{2}e^{-\frac{x^2}{2}}$ to show that the error probability, as a function of n, goes exponentially fast to zero.

Problem 4. *(Poisson Parameter Estimation)*

In this example there are two hypotheses, $H = 0$ and $H = 1$ which occur with probabilities $P_H(0) = p_0$ and $P_H(1) = 1-p_0$, respectively. The observable is $y \in \mathbb{N}_0$, i.e. y is a nonnegative integer. Under hypothesis $H = 0$, y is distributed according to a Poisson law with parameter λ_0, i.e.

$$p_{Y|H}(y|0) \quad = \quad \frac{\lambda_0^y}{y!}e^{-\lambda_0}. \tag{2}$$

Under hypothesis $H = 1$,

$$p_{Y|H}(y|1) \quad = \quad \frac{\lambda_1^y}{y!}e^{-\lambda_1}. \tag{3}$$

This example is in fact modeling the reception of photons in an optical fiber (for more details, see the Example in Section 2.2 of the notes).

1. Derive the MAP decision rule by indicating likelihood and log-likelihood ratios.
   Hint: The direction of an inequality changes if both sides are multiplied by a negative number.

2. Derive the formula for the probability of error of the MAP decision rule.

3. For $p_0 = 1/3$, $\lambda_0 = 2$ and $\lambda_1 = 10$, compute the probability of error of the MAP decision rule. You may want to use a computer program to do this.

4. Repeat (3) with $\lambda_1 = 20$ and comment.

Problem 5. *(IID versus First-Order Markov)*

Consider testing two equally likely hypotheses $H = 0$ and $H = 1$. The observable

$$Y \quad = \quad (Y_1, \dots, Y_k) \tag{4}$$

is a k-dimensional binary vector. Under $H = 0$ the components of the vector Y are independent uniform random variables (also called Bernoulli($\frac{1}{2}$) 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:

$$Pr(Y_i = y_i | Y_{i-1} = y_{i-1}, \dots, Y_1 = y_1) \quad = \quad \begin{cases} \frac{3}{4}, & \text{if } y_i = y_{i-1} \\ \frac{1}{4}, & \text{otherwise.} \end{cases} \tag{5}$$

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

2. Give a simple sufficient statistic for this decision.

3. 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

$$y \quad = \quad 0101010111111\dots111111010101\dots. \tag{6}$$

   What is the least s such that we decide for hypothesis $H = 1$? Evaluate your formula for $k = 20$.

Problem 6. *(One Bit over a Binary Channel with Memory)*

Consider communicating one bit via n uses of a binary channel with memory. The channel output Y_i at time instant i is given by

$$Y_i = X_i \oplus Z_i \qquad i = 1, \dots, n$$

where X_i is the binary channel input, Z_i is the binary noise and $\oplus$ represents modulo 2 addition. The noise sequence is generated as follows: Z_1 is generated from the distribution $\Pr(Z_1 = 1) = p$ and for $i > 1$,

$$Z_i = Z_{i-1} \oplus N_i$$

where $N_2, \dots, N_n$ are i.i.d. with $\Pr(N_i = 1) = p$. Let the codewords (the sequence of symbols sent on the channel) corresponding to message 0 and 1 be $(X_1^{(0)}, \dots, X_n^{(0)})$ and $(X_1^{(1)}, \dots, X_n^{(1)})$, respectively.

1. Consider the following operation by the receiver. The receiver creates the vector $(\hat{Y}_1, \hat{Y}_2, \ldots, \hat{Y}_n)^T$ where $\hat{Y}_1 = Y_1$ and for $i = 2, 3, \ldots, n$, $\hat{Y}_i = Y_i \oplus Y_{i-1}$. Argue that the vector created by the receiver is a sufficient statistic. Hint: Show that $(Y_1, Y_2, \ldots, Y_n)^\top$ can be reconstructed from $(\hat{Y}_1, \hat{Y}_2, \ldots, \hat{Y}_n)^\top$.

2. Write down $(\hat{Y}_1, \hat{Y}_2, \ldots, \hat{Y}_n)^\top$ for each of the hypotheses. Notice the similarity with the problem of communicating one bit via n uses of a binary symmetric channel.

3. How should the receiver choose the codewords $(X_1^{(0)}, \ldots, X_n^{(0)})$ and $(X_1^{(1)}, \ldots, X_n^{(1)})$ so as to minimize the probability of error? Hint: When communicating one bit via n uses of a binary symmetric channel, the probability of error is minimized by choosing two codewords that differ in each component.