ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE
School of Computer and Communication Sciences

Principles of Digital Communications: Summer Semester 2012

Assignment date: April 4th, 2012, 3:15pm
Due date: April 4th, 2012, 5:15pm

Solution to Midterm Exam

You are allowed one A4-sized piece of paper as crypt sheet. No course book, exercises, or electronic devices are allowed. There are four problems. We do not presume that you will finish all of them. Choose the ones you find easiest and collect as many points as possible. Good luck!

Name: ______________________________

Problem I	/ 35
Problem II	/ 30
Problem III	/ 25
Problem IV	/ 25
Total	/115

Problem 1. *(Yes/No Questions)* [35pts] – Give one-line justifications.

1. [5pts] Assume that we have a hypothesis testing problem with the observation (Y_1, Y_2). Then $(Y_1, Y_1 + Y_2)$ is a sufficient statistic. [YES / NO]

   YES. Because the transformation is one to one and we can completely recover (Y_1, Y_2) from $(Y_1, Y_1 + Y_2)$ so we don't lose any information.

2. [5pts] Let X and Y be independent and distributed according to $\mathcal{N}(0, \sigma^2)$. Let $X = R\cos(\Phi), Y = R\sin(\Phi)$ be the polar transformation. Then R and Φ are independent. [YES / NO]

   YES. After transformation we obtain that $f_{R,\Phi}(r, \phi) = \frac{1}{2\pi}\frac{r}{\sigma^2}e^{-\frac{r^2}{2\sigma^2}}$ which can be written in the factorized form. This show that R and Φ are independent.

3. [5pts] Let $X \sim \mathcal{N}(0, \sigma^2)$ and $Y \sim \mathcal{N}(0, \sigma^2)$. Then $X + Y \sim \mathcal{N}(0, 2\sigma^2)$. [YES / NO]

   NO. This is not true in general. For example assume that $X \sim \mathcal{N}(0, \sigma^2)$ and $Y = -X$ then it is easy to show that $X \sim \mathcal{N}(0, \sigma^2)$ but $X + Y = 0$ which has variance 0 and hence is not $\mathcal{N}(0, 2\sigma^2)$. But if we have the independence assumption then it is easy to show that $X + Y \sim \mathcal{N}(0, 2\sigma^2)$.

4. [5pts] If (X_1, X_2) are jointly Gaussian random variables with a covariance matrix $\Sigma = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}$ then the variance of $X_1 + X_2$ is 3. [YES / NO]

   YES. It is easy to show that

$$\begin{aligned} \mathrm{var}(X_1 + X_2) &= \mathrm{var}(X_1) + \mathrm{var}(X_2) + 2\,\mathrm{cov}(X_1, X_2) \\ &= 1 + 1 + 2 \times 0.5 = 3. \end{aligned}$$

5. [5pts] Assume $\phi = (1, 0, 0)$ is a three dimensional vector. Let $\psi_1 = (\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0)$, $\psi_2 = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0)$ and $\psi_3 = (0, 0, 1)$ be an orthonormal basis for the three dimensional space and let $(\alpha_1, \alpha_2, \alpha_3)$ be the representation of ϕ in this basis. In other words, $\phi = \alpha_1\psi_1 + \alpha_2\psi_2 + \alpha_3\psi_3$. Then $\sqrt{\alpha_1^2 + \alpha_2^2 + \alpha_3^2} = \sqrt{2}$. [YES / NO]

NO. We know that the norm and the inner product are invariant under coordinate transformations. In the initial coordinate we have $\|\phi\| = \sqrt{1+0+0} = 1$. Hence the $\sqrt{\alpha_1^2 + \alpha_2^2 + \alpha_3^2} = \|\phi\| = 1$.

6. [5pts] Assume that we have a binary hypothesis testing problem with observation in $\mathbb{R}^n$. Let $\mathcal{R}_0$ and $\mathcal{R}_1$ be the MAP decision regions where $\mathcal{R}_0 \cup \mathcal{R}_1 = \mathbb{R}^n$. If we enlarge the region $\mathcal{R}_1$ then (increasing here means *non-decreasing*, i.e., constant or strictly increasing):

   (i) [2.5pts] The conditional error probability under H_1 increases. [YES / NO]

   NO. As we make $\mathcal{R}_1$ larger $\mathcal{R}_0$ becomes smaller. Hence, by definition, the conditional error probability under H_1, $Pr\{E|H_1\} = \int_{\mathcal{R}_0} f_{Y|H_1}(y)\, dy$, decreases.

   (ii) [2.5pts] The average error probability increases. [YES / NO]

   YES. In general if we change the decision regions then our decision rule is not optimal and hence its average error probability is higher than the MAP rule. Notice that the MAP rule is the decision rule minimizing the **average** error probability.

7. [5pts] For a binary hypothesis testing problem the likelihood ratio is a sufficient statistic. [YES / NO]

   YES. You have proved this in the Homework 4, Problem 5, Part 3.

Problem 2. *(Scaling Law)* [30pts]

We want to send one bit of information. To transmit 0, we send the vector $X = (-A, -A, \ldots, -A)$ of length n, where A is a given amplitude. To transmit 1, we send the vector $X = (A, A, \ldots, A)$. The observation is $Y = X + Z$, where $Z = (Z_1, Z_2, \ldots, Z_n)$ is the additive noise of the channel consisting of iid components distributed like $\mathcal{N}(0, \sigma^2)$ where $\sigma = 1$. Assume that 0 and 1 are equiprobable.

1. [10pts] Derive the MAP rule and find the error probability as a function of n, A, and σ.

   MAP rule:

   $$\sum_{i=1}^{n} Y_i \overset{H=1}{>}\overset{H=0}{<} 0.$$

   Error probability:

   $$P_e = Q(\frac{d}{2\sigma}) = Q(\frac{\sqrt{n}A}{\sigma}).$$

2. [5pts] Suppose that the noise of the channel increases to $\sigma = 2$. Assume that n is fixed but we can change A. Call the new value A'. How shall we set A' so that the error probability stays the same.

   $A' = 2A$

3. [5pts] Suppose that the noise of the channel increases to $\sigma = 2$. Assume that we can't change A but we can increase n. Call the new value n'. How shall we choose n' so that the error probability stays the same.

   $n' = 4n$

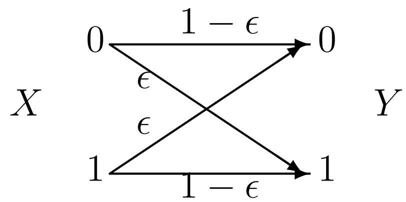
Now consider a general m-ary Gaussian hypothesis testing problem. The observation under hypothesis i is $Y = s_i + Z$, where $s_i \in \mathbb{R}^n$, and where Z is vector of n iid components distributed like $\mathcal{N}(0, \sigma^2)$. Assume that all priors are equal and let P_e denote the average error probability under the MAP decision rule.

4. [10pts] Suppose that σ is very small (and hence P_e is also *very small*). Assume now that σ is changed to $\sigma' = \sigma/\sqrt{2}$. Let the new error probability be P_e'. Expressed in terms of P_e, roughly, how large is P_e'? (**Hint:** You may find the following bound useful, $Q(x) \leq \frac{1}{2}e^{-\frac{x^2}{2}} \quad x \geq 0$).

   We know that σ is very small. This implies that the error probability is roughly equal to $Q(d_{\min}/(2\sigma))$, where $d_{\min}$ is the smallest Euclidean distance between two neighboring points. This is true since we know that the Q-function decays rapidly as function of its argument. Further, $Q(x)$ is upper bounded by $\frac{1}{2}e^{-\frac{x^2}{2}}$. Hence, the error probability can be approximated by $e^{-\frac{d_{\min}^2}{8\sigma^2}}$. Therefore, if we decrease the noise variance by a factor $\sqrt{2}$ the value of the exponent is doubled which implies that the error probability is essentially squared. In other words, $P_e' \sim P_e^2$.

Problem 3. *(Binary Symmetric Channel)* [25pts]

Let the communication channel between a transmitter and a receiver be a BSC (Binary Symmetric Channel) with input X, output Y, and bit flip probability $\boldsymbol{\epsilon < \frac{1}{2}}$. In other words, as shown in the figure below, the input and output of the channel is 0 or 1, $P(Y = 0|X = 0) = P(Y = 1|X = 1) = 1 - \epsilon$ and $P(Y = 1|X = 0) = P(Y = 0|X = 1) = \epsilon$.



Assume that $p = P(X = 1)$. In this problem we are going to investigate the performance as a function of p.

1. [5pts] Let $L(Y) = \frac{P(Y|X=1)}{P(Y|X=0)}$ denote the likelihood ratio. Express $L(0)$ and $L(1)$ as a function of ϵ.

$$L(0) = \frac{P(Y = 0 \mid X = 1)}{P(Y = 0 \mid X = 0)} = \frac{\epsilon}{1 - \epsilon},$$
$$L(1) = \frac{P(Y = 1 \mid X = 1)}{P(Y = 1 \mid X = 0)} = \frac{1 - \epsilon}{\epsilon}.$$

2. [10pts] Derive the MAP decision rule and find the error probability for the following ranges of p. (**Hint:** In every case, simply try to find the decision result for two possible observations $Y = 0$ and $Y = 1$ and also use the fact that $\epsilon < \frac{1}{2}$.)

   (a) $p < \epsilon$.

   (b) $p \in [\epsilon, 1 - \epsilon]$.

   (c) $p > 1 - \epsilon$.

   In this problem our observation can take just two values $0, 1$. Hence, it is sufficient two find the decision result for these two cases. We know that for the MAP decision rule, one must obtain the likelihood ratio as a function of the observation and compare it with the threshold $\frac{p_0}{p_1}$ to obtain the decision result. As we change the prior probability p this threshold changes and our decision changes as well. Let us discuss the three cases:

   (a) $p < \epsilon$: In this case one can simply check that $L(0)$ and $L(1)$ are both less than the threshold $\frac{p_0}{p_1} = \frac{1-p}{p}$. Assume we observe $Y = 0$ at the output of the channel then we have to compare $L(0)$ with the threshold. As the threshold is larger than $L(0)$

we have to decide on 0. Similarly if we observe $Y = 1$ at the output of the channel we have to compare $L(1)$ with the threshold. But again the threshold is higher than $L(1)$ hence according to the MAP rule we have to decide on 0. In other words, we can simply say that $D(Y) = 0$ where D denotes the decision result and notice that the decision result is a function of the observation Y, the output of the channel. We also have $Pr\{E|H_0\} = 0$ and $Pr\{E|H_1\} = 1$ because we always decide on 0 and if the hypothesis is H_0 we don't make any error whereas under H_1 hypothesis we have complete error because according to the decision rule we never decide on 1. We can simply obtain

$$\begin{aligned} Pr\{E\} &= p_0 \times 0 + p_1 \times 1 \\ &= p_1 = p. \end{aligned}$$

(b) $\epsilon < p < 1 - \epsilon$: In this case one can obtain that $D(Y) = Y$. In other words if the observation is 0 then the decision result is 0 and if the observation is 1 then the decision result is 1. You can simply check this by comparing $L(0)$ and $L(1)$ with the decision threshold $T \triangleq \frac{p_0}{p_1}$ and prove that in this case $L(0) < T$ and $L(1) > T$.

$$\begin{aligned} Pr\{E|H_0\} &= Pr\{D(Y) = 1|H_0\} \\ &= Pr\{Y = 1|H_0\} \\ &= Pr\{Y = 1|X = 0\} = \epsilon. \end{aligned}$$

Similarly one can obtain that $Pr\{E|H_1\} = \epsilon$. Hence we have

$$Pr\{E\} = p_0 \times \epsilon + p_1 \times \epsilon = \epsilon.$$

(c) $p > 1 - \epsilon$: In this case one can show that $D(Y) = 1$. In other word, wether we observe 0 or 1 we decide on 1. It can be also shown that $Pr\{E\} = 1 - p$

3. [5pts] Draw the curve of the error probability as a function of p and specify the important points on the curve.

   The curve first rises linearly until $p = \epsilon$, then stays constant until $1 - \epsilon$, and then decreases again linearly until 1. A typical shape is shown in Figure (1) for $\epsilon = 0.3$.

4. [5pts] What is the worst range of values of p that results in the maximum error probability and what is this maximum error probability?

   One can simply check that the worst range is $p \in [\epsilon, 1 - \epsilon]$ resulting in an error probability of ϵ.

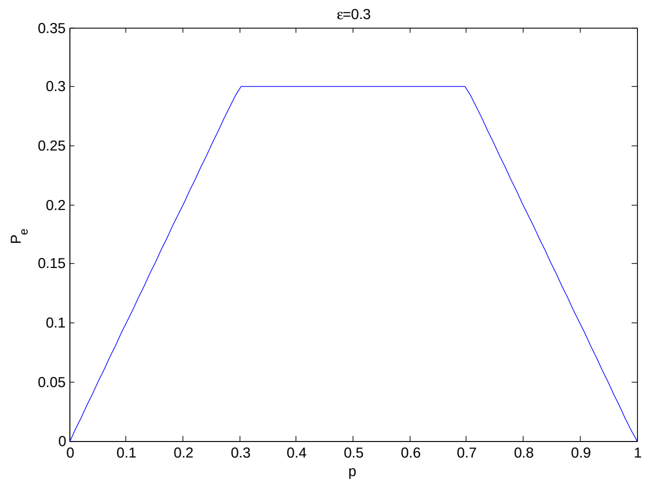


Figure 1: MAP error as a function of prior probability

Problem 4. *(Simple Wireless Channel)* [25pts]

A transmitter-receiver pair uses a wireless channel to communicate. When the transmitter sends a pulse of amplitude A, neglecting the effect of noise, what is received is a zero-mean Gaussian random variable of variance A^2. In addition, the received signal is perturbed by additive noise $\mathcal{N}(0, \sigma^2)$ which is independent of the transmitted signal. The transmitter uses OOK (On-Off Keying) signaling. More precisely, to send 1, the transmitter sends a signal of amplitude A, and to send 0 the transmitter is off and doesn't send anything. We assume that 0 and 1 are equiprobable. We denote the received signal by R.

1. [10pts] Formulate the problem as a binary hypothesis testing problem and find the probability density function of the observation R under two hypotheses.

   We have:

$$R \sim \mathcal{N}(0, \sigma^2), \;\; H = 0,$$
$$R \sim \mathcal{N}(0, A^2 + \sigma^2), \;\; H = 1.$$

2. [10pts] Derive the MAP rule and find its error probability as a function of $\text{SNR} \triangleq \frac{A^2}{\sigma^2}$.

$$\frac{1}{\sqrt{2\pi(\sigma^2+A^2)}}e^{-\frac{y^2}{2(\sigma^2+A^2)}} \overset{H=1}{\geq} \overset{H=0}{\leq} \frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{y^2}{2\sigma^2}},$$

$$e^{\frac{y^2}{2\sigma^2}(1-\frac{1}{\beta})} \overset{H=1}{\geq} \overset{H=0}{\leq} \sqrt{\beta}, \quad \beta = 1+\frac{A^2}{\sigma^2},$$

$$|y| \overset{H=1}{\geq} \overset{H=0}{\leq} \sigma\sqrt{\frac{\log(\beta)}{(1-\frac{1}{\beta})}} = \gamma.$$

We can also obtain that

$$\begin{aligned} P_e &= \frac{1}{2}(2\,Q(\frac{\gamma}{\sigma})) + \frac{1}{2}(1-2Q(\frac{\gamma}{\sqrt{\sigma^2+A^2}}), \\ &= Q(\sqrt{\frac{\log(\beta)}{1-\frac{1}{\beta}}}) + \frac{1}{2}(1-2Q(\sqrt{\frac{\log(\beta)}{\beta-1}})). \end{aligned}$$

3. [5pts] Show that the error probability goes to zero as SNR tends to infinity.

   Note that β tends to infinity. Using $\lim_{x\to\infty} Q(x) = 0$ and $Q(0) = \frac{1}{2}$ we see that both terms tend to 0.