EE226: Random Processes in Systems **Fall'06**

Problem Set 3 — Due Oct, 5

Lecturer: Jean C. Walrand *GSI: Assane Gueye*

This problem set essentially reviews detection theory. Not all exercises are to be turned in. Only those with the sign ★ are due on *Thursday, October* 5^{th} at the beginning of the class. Although the remaining exercises are not graded, you are encouraged to go through them.
We will discuss some of the exercises during discussion sections.
Please feel free to point out errors and notions that need to be clarified.

Exercise 3.1. ★
We have seen in class that a convenient way to write the error probability in Gaussian detection is the use of the Q-function defined as:

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} exp\{-\frac{t^2}{2}\}dt$$

In this exercise we will study some properties of the Q-function:
(a) Show that $Q(-x) = 1 - Q(x)$
(b) Show that

$$\left(1 - \frac{1}{x^2}\right) \frac{e^{-\frac{x^2}{2}}}{x\sqrt{2\pi}} \leq Q(x) \leq \frac{e^{-\frac{x^2}{2}}}{x\sqrt{2\pi}}$$

Hint: rewrite the integrand (multiply and divide by t) and use integration by parts.

Solution:
(a)We will use the definition and a change of variable.

$$\begin{aligned}
Q(-x) = \frac{1}{\sqrt{2\pi}} \int_{-x}^{\infty} e^{-\frac{t^2}{2}} dt &= \frac{1}{\sqrt{2\pi}} \int_{-x}^{0} e^{-\frac{t^2}{2}} dt + \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} e^{-\frac{t^2}{2}} dt \\
\text{change of variable } x \leftrightarrow -x &= \frac{1}{\sqrt{2\pi}} \int_0^{x} e^{-\frac{t^2}{2}} dt + \frac{1}{2} \\
&= \frac{1}{2} - \frac{1}{\sqrt{2\pi}} \int_x^{+\infty} e^{-\frac{t^2}{2}} dt + \frac{1}{2} \\
&= 1 - Q(x)
\end{aligned}$$

(b)For the upper and lower bounds we will use integration by parts. We will drop the $\sqrt{2\pi}$ term in the calculations.

$$\begin{aligned}
\int_x^\infty e^{-\frac{t^2}{2}} &= \int_x^\infty \frac{t}{t} e^{-\frac{t^2}{2}} \\
(u = t, v' = te^{-\frac{t^2}{2}}) &= \left[-\frac{e^{-\frac{t^2}{2}}}{t} \right]_x^\infty - \int_x^{+\infty} \frac{1}{t^2} e^{-\frac{t^2}{2}} dt \qquad (3.1) \\
(\text{second term} \geq 0) &\geq \frac{e^{-\frac{x^2}{2}}}{x}
\end{aligned}$$

For the upper bound we will perform another integration by parts in the second term of equation 3.1.

$$\begin{aligned}
\int_x^\infty e^{-\frac{t^2}{2}} &= \frac{e^{-\frac{x^2}{2}}}{x} - \int_x^{+\infty} \frac{t}{t \cdot t^2} e^{-\frac{t^2}{2}} dt \\
(u = \frac{1}{t^3}, v' = te^{-\frac{t^2}{2}}) &= \frac{e^{-\frac{x^2}{2}}}{x} - \left[-\frac{e^{-\frac{t^2}{2}}}{t^3} \right]_x^\infty + \int_x^{+\infty} \frac{1}{t^4} e^{-\frac{t^2}{2}} dt \\
(3^{rd} \text{ term} \geq 0) &\geq \frac{e^{-\frac{x^2}{2}}}{x} - \frac{e^{-\frac{x^2}{2}}}{x^3}
\end{aligned}$$

Exercise 3.2. *The output of a system is normal $\mathcal{N}(1.1, 1.1)$ if the input X is equal to 1. When the input X is equal to 2, the output is normal $\mathcal{N}(0.1, 0.1)$.*
Observing the output of the system, we would like to compute our best estimate of X_i (or estimate of i).
Find the best estimate of $i \in 1, 2$ given the output of the system (best in term of maximizing the probability of X_i given the system output).
We assume that the input is X_1 with probability 0.6.

Solution:(Hint)
The solution is given in the 126 notes (Example 8.6.1)

Exercise 3.3. *The Binary Phase-shift Keying (BPSK) and Quadrature Phase-shift Keying (QPSK) modulation schemes are shown in figure 3.1. We consider that in both cases, the symbols (S) are sent over an additive gaussian channel with zero mean and variance σ^2. Assuming that the symbols are equally likely, compute the average error probability for each scheme. Which one is better?*

Solution:(Hint)
Note that the comparison is not fair because the two schemes do not have the same rate (eggs and apples!). But let us compute the error probabilities and compare them.
The BPSK error probability is the one given in example 3.1 of Gallager's notes (ML detection with $b(gal) \leftrightarrow a$ and $a(gal) \leftrightarrow -a$). Thus the error probability is given by

$$P_e^{bpsk} = Q(\frac{2a}{\sigma})$$

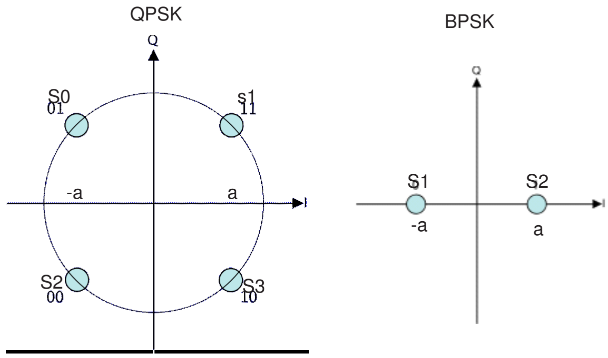

Figure 3.1. BPSK and QPSK modulations

For QPSK, consider that signal S_1 was sent and observe that error occurs if the received signal does not fall in the first quartan. By considering the probability of detecting any other signal, we can see that the error probability is equal to

$$P_e^{qpsk} = Q(\frac{2\sqrt{2}a}{\sigma}) + 2Q(\frac{2a}{\sigma})$$

For a large enough we have $P_e^{qpsk} \approx 2P_e^{bpsk}$

Exercise 3.4. *Given $X = i, (i = 1, 2)$, Z is an exponential random variable with mean $\frac{1}{\lambda_i}$. Give the MAP decision rule of X given the observation Z. What is the probability of error? (For the MAP assume that $X = 1$ with probability p).*

Solution:(Hint)
Solution given in 126 notes example 8.6.2

Exercise 3.5. *The conditional pmf of Y given X is given by the table below.*

$X \diagdown Y$	$+1$	-1
$+1$	0.3	0.7
-1	0.6	0.4

Find the MAP and MLE of X given Y.
(For the MAP assume that $X = 1$ with probability p).

Solution:(Hint)
Solution given in 126 notes example 8.6.7

Exercise 3.6. *The purpose of a radar system is to detect the presence of a target and extract useful information about the target. Suppose that in such a system, hypothesis H_0 is that there is no target present, so that the observation is $s = w$, where w is a normal random variable with zero mean and variance σ^2. Hypothesis H_1 corresponds to the presence of the*

target and the observation is $s = a + w$ where a is an echo produced by the target and is assumed to be known. Evaluate:
(1) The probability of false alarm *defined as the probability that the receiver decides a target is present when it is not.*
(2) The probability of miss detection *defined as the probability that the receiver decides a target is not present when it is.*

Solution:(Hint)
Write down the decision rule: H_0 if $y < a/2$ and H_1 otherwise and compute the errors.

Exercise 3.7. ★
Suppose that we want to send at the same time two independent binary symbols $X_1 \in \{-1, 1\}$ and $X_2 \in \{-1, 1\}$ over an additive Gaussian channel, where -1 and $+1$ are equally likely. A clever communication engineer proposes a transmission scheme that uses twice the channel to send at each time a linear combination of X_1 and X_2. The channel input output relation is given by:

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} h_1 & h_2 \\ -h_2 & h_1 \end{pmatrix} * \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} + \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix}$$

where Y_1 and Y_2 are the channel outputs, the $h_i, i = 1, 2$ are known real constant, and $Z_i, i = 1, 2$ are iid standard normal random variables.
We are interested in detecting both X_1 and X_2 (i.e. the vector $(X_1, X_2)^T$).
(a) Show that the vector *detection problem can be decomposed into two scalar detection problems.*
Hint: observe that the rows of the matrix H are orthogonal.
(b) Give the decision rule for detecting X_1 and write the error probability using the Q-function.

Solution:
The scheme presented in this exercise in the Alamouti scheme well known in communication systems (mainly in multi-antenna communications).
(a) Observing that the rows (and columns) of the matrix $\mathbf{H}$ are orthogonal, let us project the received signal on the directions $\mathbf{H_1} = (h_2, h_1)^T$ and $\mathbf{H_2} = (h_1, -h_2)^T$.
We have:

$$Y_2 = \mathbf{H_1}^T Y = (0, h_1^2 + h_2^2) Y = (h_1^2 + h_2^2) x_2 + h_2 z_1 + h_1 z_2$$
$$Y_1 = \mathbf{H_2}^T Y = (h_1^2 + h_2^2, 0) Y = (h_1^2 + h_2^2) x_1 + h_1 z_1 - h_2 z_2$$

Thus we have written the vector detection problem into two scalar detection problems $Y_i = \alpha x_i + w_i, i = 1, 2$ where $w_i \sim \mathcal{N}(0, h_1^2 + h_2^2)$. To be sure that the problem can be decoupled, we need to check that the w_i's are uncorrelated.

$$\begin{aligned}
[E[w_1 w_2] &= E[(h_2 z_1 + h_1 z_2)(h_1 z_1 - h2_z 2)] \\
&= h_1 h_2 E[z_1^2] - h_2^2 E[z_1 z_2] + h_1^2 E[z_1 z_2] - h_1 h_2 E[z_2^2] \\
&= h_1 h_2 - 0 + 0 - h_1 h_2 \\
&= 0
\end{aligned}$$

Thus w_1 and w_2 are uncorrelated. (Since they are gaussian, they are also independent). Hence the vector detection problem can be decoupled into 2 independent scalar detection problems.

(b) We can use the same technique as in example 3.1 of the lecture notes. We decide $\hat{x}_1 = -1$ if $y_1 < 0$ and $\hat{x}_1 = 1$ if $y_1 > 0$. Since the two symbols are equally likely, the error probability is equal to the probability of error given that $x_1 = -1$ was sent. Given $x_1 = -1$ was sent, $Y_1 \sim \mathcal{N}(-(h_1^2 + h_2^2), h_1^2 + h_2^2)$

$$\begin{aligned}
P_e &= P(N(-(h_1^2 + h_2^2), h_1^2 + h_2^2) > 0) \\
&= P(N(0,1) > \sqrt{h_1^2 + h_2^2}) \\
&= Q\left(\sqrt{h_1^2 + h_2^2}\right)
\end{aligned}$$

Exercise 3.8. *In this exercise we consider a binary detection problem (like in example 3.1) where n independent observations $Y_i = a + Z_i$ are available. Assume that the $Z_i, i = 1, \ldots, n$ are iid $\mathcal{N}(0, \sigma^2)$. Hypothesis H_0 and H_1 are mapped to $a = -1$ and $a = +1$ respectively. Find the decision rule for the ML and the MAP detectors.*
Hint: find a sufficient statistic first.

Solution:(Hint)
Convince yourself that the sample mean $\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$ is a sufficient statistic and that the decision rule is given by: H_0 if $\bar{Y} < 0$ and H_1 otherwise.
Now compute the distribution of $\bar{Y}$ and deduce the error probabilities.

Exercise 3.9. *We observe n independent realizations $(Y_1, \ldots, Y_n)$ of a random variable Y which distribution depends on another variable X. If $X = 1$ (resp. 2), then Y is an exponential random variable with mean 1 (resp. 2). We assume that X takes the value 1 with probability $p \in (0, 1)$.*
Compute the optimal detector of X given the Y_i's (optimal in terms of maximizing the probability of X given the observation $Y_1^n = (Y_1, \ldots, Y_n)$).
What is this detector when $p = 1/2$.

Solution:(Hint)
Solution given in example 8.6.8 of 126 notes.

Exercise 3.10. ★

Consider a modified version of Example 3.2 of the course note (Gallager) where the source output is no longer 0 or 1 but takes values in the set $S = \{1, 2, \ldots, I\}$. *If the source output is* $i \in S$, *the modulator produces the real vector* $\vec{a_i} = (a_{i1}, \ldots, a_{in})$. *Given that the source output is* i *the observation is* $\vec{Y} = \vec{a}_i + \vec{Z}$ *where the noise is assumed to be* $\mathcal{N}(0, I_n)$.

(a) Show that the ML decision rule is given by

$$\hat{i} = argmin_i(\| \vec{Y} - \vec{a}_i \|^2)$$

where $\| \,.\, \|$ *is the Euclidean distance in* R^n.

(b) For $n = 2$, *use a one dimensional sufficient statistics to compute the decision rule.*

(c) Compare your decision rule to the one in (a). Explain.

Solution:

(a) The ML decision point is the one that maximizes the likelihood (or conditional distribution). It is given by

$$\begin{aligned}
\hat{i} &= argmax_i f_{Y|H}(y|i) \\
&= argmax_i\left(\frac{1}{(2\pi)^{\frac{n}{2}}}\exp(-\frac{1}{2}(Y-a_i)(Y-a_i)^T)\right), \text{By def. of the cond. dist} \\
&= argmax_i(-\frac{1}{2}(Y-a_i)(Y-a_i)^T)), \text{Taking log...} \\
&= argmin_i(Y-a_i)(Y-a_i)^T) \\
&= argmin_i\|Y-a_i\|
\end{aligned}$$

Thus the decision point is the one that minimizes the euclidian distance between the receive signal Y and the data symbols a_i.

(b) A one-dimensional sufficient statistic (when $n = 2$) is $\Theta = (a_2 - a_1)^T Y$. This is the inner product between Y and the direction of the signal.

Given that $S_i, i = 1, 2$ was sent, $\Theta \sim \mathcal{N}((a_2 - a_1)^T a_i, \|a_2 - a_1\|^2)$. Now we can compute the $LLR(\Theta)$ which is equal to

$$\begin{aligned}
LLR(\Theta) &= 2\Theta - (a_2 - a_1)^T(a_2 + a_1) \\
&= 2(a_2 - a_1)^T y - (a_2 - a_1)^T(a_2 + a_1) \\
&= 2(a_2 - a_1)^T(y - \frac{(a_2 + a_1)}{2})
\end{aligned}$$

We will then decide that $\hat{S} = S_2$ whenever $LLR(\Theta) \geq 0$.

To understand this decision rule, note first $y - \frac{(a_2+a_1)}{2}$ is just changing the origin from $(0, 0)$ vector to $\frac{(a_2+a_1)}{2}$ (to simplify we assume that $\|a_2 - a_1\| = 1$). Think of the axes to be changed also from the usual ones to $(a_2 - a_1)$ (the direction of the signal) and $(a_2 - a_1)^\perp$ (the direction perpendicular to the signal). Let $y = \alpha(a_2 - a_1) + \beta(a_2 - a_1)^\perp$ in the new coordinate system.

Taking the inner product

$$\begin{aligned}(a_2-a_1)^T y &= (a_2-a_1)^T(\alpha(a_2-a_1)+\beta(a_2-a_1)^{\perp})\\ &= \alpha(a_2-a_1)^T(a_2-a_1)+\beta(a_2-a_1)^T(a_2-a_1)^{\perp}\\ &= \alpha\end{aligned}$$

Our decision rule becomes the simple test of whether α is greater then $\frac{(a_2-a_1)^T(a_2+a_1)}{2}$. Geometrically, we have reduced our vector detection problem to a one-dimensional detection problem where the detection is along the direction of the signal.

Important remarks

(1) Our decomposition was possible because the noise was gaussian, uncorrelated and unit variance.

(2) If the noise is still gaussian but correlated with arbitrary (non singular) covariance matrix, we need to re-scale the received signal before the projection. This is the job of the term K^{-1} in the general sufficient statistic $(a_2-a_1)K_Z^{-1}Y$.

(3) The case where the noise is gaussian with singular covariance matrix is discussed in Problem 3.12

(4) For a general distribution of the noise, there is no such result (as far as I know!)...This gives an idea why Gaussian's are important for us.

(c) It is not hard to verify that the two decision rules are the same (they have to be!). For that, note that in part (a) we decide that $S=S_1$ if

$$\begin{aligned}& \|y-a_1\|^2 < \|y-a_2\|^2\\ \Leftrightarrow\ & \|y\|^{-}2a_1^T y+\|a_1\|^2 < \|y\|^{-}2a_2^T y+\|a_2\|^2\\ \Leftrightarrow\ & 2(a_2-a_1)^T y-(\|a_2\|^2-\|a_1\|^2)<0\\ \Leftrightarrow\ & 2(a_2-a_1)^T y-(a_2-a_1)^T(a_2+a_1)<0\\ \Leftrightarrow\ & (a_2-a_1)^T(y-\frac{a_2+a_1}{2}<0\end{aligned}$$

which is the decision in part (b).

The method in part (b) is quite intuitive and brings us back to the simple scalar detection case but it is applicable only for binary detection. If there are more than 2 hypothesis, only the method in part (a) can be used.

Exercise 3.11. ★

Gallager's note: Exercise 3.1

Solution:(Hint)

This exercise was just verification of results presented in the Galalger's note.

(a) Given H_0, an error occurs if $LLR(\overrightarrow{Y})\geq 0$. Write the error probability to get 3.31.

(b) Similar to part (a).

(c)

$$E[U]=E[(\overrightarrow{b}-\overrightarrow{a})^T\overrightarrow{Y}]=(\overrightarrow{b}-\overrightarrow{a})^T E[\overrightarrow{Y}]$$

Thus given H_0 $E[U] = (\overrightarrow{b} - \overrightarrow{a})^T \overrightarrow{a}$ and given H_1 $E[U] = (\overrightarrow{b} - \overrightarrow{a})^T \overrightarrow{b}$
The variances are given by:

$$var(U|H_0) = \sigma^2 \| \overrightarrow{b} - \overrightarrow{a} \|^2 = var(U|H_0)$$

Now using the decision rule given by 3.27, we easily get 3.31.
(d) Again U is a Gaussian random variable with mean and variance given above. Just write the ratio of the conditional distribution and take the log.
(e) Since U is one dimensional, we can now apply 3.18 and get 3.31 again.

The whole point of this exercise was to study properties of the sufficient statistic. The message to be taken is that the sufficient statistic is easier to work with (since it is usually of lower dimension than the original variable), it has the same LLR and leads to the same decision rule as the observation.

Exercise 3.12. ★
Gallager's note: Exercise 3.2

Solution:(Hint)
Parts (a), (b), (c) are very similar to what we did in the previous problem. The only tricky part is (d).
(d) As it has been noticed in previous homework, if the correlation matrix K_Z is singular, then the noise lies in a lower dimensional space. If $\overrightarrow{b} - \overrightarrow{a}$ is entirely in that space, then the problem can be reduced to a lower dimensional estimation problem. If some component of $\overrightarrow{b} - \overrightarrow{a}$ is not in the space, then that component will be received without noise. By observing that particular component, we can always detect without error.

Exercise 3.13. ★
Gallager's note: Exercise 3.5

Solution:(Hint)
(a) Observe that given H_0, $Y \sim \mathcal{N}(a, AAT)$ and given H_1, $Y \sim \mathcal{N}(b, AA^T)$.
The log-likelihood ratio is then given by:

$$\begin{aligned} LLR(Y) &= -\frac{1}{2}\left[2(b-a)^T(A^{-1})^T A^{-1} Y + a(AA^T)^{-1}a - b(AA^T)^{-1}b\right] \\ &= -\frac{1}{2}\left[2(b-a)^T(A^{-1})^T v + a(AA^T)^{-1}a - b(AA^T)^{-1}b\right] \end{aligned}$$

which shows that the dependency of the LLR to Y, is only through $A^{-1}Y = v$.
Thus v is a sufficient statistic.
(b) Observe that given H_0, $v \sim \mathcal{N}(A^{-1}a, I)$ and given H_1, $v \sim \mathcal{N}(A^{-1}b, I)$. Now compute the LLR of verify that it is equal to the one computed in (a).
(c) Note that the noise is uncorrelated and we can apply the techniques used in example 3.2

of Gallager's notes.
After some calculation steps, we obtain:

$$Pr(e|H_0) = Q(\frac{log(\eta)}{\gamma} + \frac{\gamma}{2}) \quad \text{where} \quad \gamma = \|A^{-1}(b-a)\|$$

Note: The goal of this exercise was to study the whitening filter. Whenever the noise covariance matrix can be written as $K_Z = AA^T$ with A a nonsingular matrix (this is always the case if K_Z is not singular), then the detection problem $Y = a_i + Z$ is equivalent to $\tilde{Y} = A^{-1}a_i + W$ where $\tilde{Y} = A^{-1}Y$ and W is a zero mean, uncorrelated ($\mathcal{N}(0, I)$ in case of gaussian) noise.
The process of multiplying by A^{-1} is called the *whitening filter*. You will see more of it later in the course.

Exercise 3.14. ★
Gallager's note: Exercise 3.9

Solution:
Conditioned on $H = 0$, $(Y1, Y2)$ and $(Y3, Y4)$ are independent. Each of these two-dimensional distributions is circularly symmetric in the plane, so the conditional density of $(Y1, Y2, Y3, Y4)$ given $H = 0$ must depend on $(Y1, Y2, Y3, Y4)$ only through $Y_1^2 + Y_2^2$, and $Y_3^2 + Y_4^2$. Since this is also true for the conditional density given $H = 1$, $(Y_1^2 + Y_2^2, Y_3^2 + Y_4^2) = (V_0, V_1)$ must be sufficient statistic. By the symmetry of the transmitted signal and the noise,

$$f_{V_0,V_1|H}(v_0, v_1|1) = f_{V_0,V_1|H}(v_1 v_0|0)$$

We also expect that $f_{V_0,V_1|H}(v_0, v_1|0) > f_{V_0,V_1|H}(v_1, v_0|0)$ whenever $v_0 > v_1$ (proving this is technical, and not required). Then the decision rule must be the one stated.

Technical stuff follows:

$$\begin{aligned}
f_{Y|H}(y|0) &= \int_0^{2\pi} f_{Y,\phi|H}(y,\theta|0)d\theta \\
&= \int_0^{2\pi} f_{Y|H}(y|0)f_\phi(\theta)d\theta \\
&= \int_0^{2\pi} \frac{1}{2\pi}\frac{1}{(2\pi)^2\sigma^4}\exp\left(-\frac{(y_1-acos\theta)^2+(y_2-asin\theta)^2+y_3^2+y_4^2}{2\sigma^2}\right)d\theta \\
&= \frac{1}{(2\pi)^2\sigma^4}\exp\left(-\frac{y_1^2+y_2^2+y_3^2+y_4^2+a}{2\sigma^2}\right)\int_0^{2\pi}\frac{1}{2\pi}\exp\left(a(y_1cos\theta+y_2sin\theta)/\sigma^2\right)d\theta \\
&= \frac{1}{(2\pi)^2\sigma^4}\exp\left(-\frac{y_1^2+y_2^2+y_3^2+y_4^2+a}{2\sigma^2}\right) \\
&\quad \cdot\int_0^{2\pi}\frac{1}{2\pi}\exp\left(\frac{a\sqrt{y_1^2+y_2^2}}{\sigma^2}\cos(\theta-\arctan(y_2/y_1))\right)d\theta \\
&= \frac{1}{(2\pi)^2\sigma^4}\exp\left(-\frac{y_1^2+y_2^2+y_3^2+y_4^2+a}{2\sigma^2}\right)\int_0^{2\pi}\frac{1}{2\pi}\exp\left(\frac{a\sqrt{y_1^2+y_2^2}}{\sigma^2}\cos(\psi)\right)d\psi \\
&= \frac{1}{(2\pi)^2\sigma^4}\exp\left(-\frac{y_1^2+y_2^2+y_3^2+y_4^2+a}{2\sigma^2}\right)I_0(\frac{a\sqrt{y_1^2+y_2^2}}{\sigma^2})
\end{aligned}$$

where I_0 is the modified Bessel function of the first kind and $0-$th order. Similarly,

$$f_{Y|H}(y|1) = \frac{1}{(2\pi)^2\sigma^4}\exp\left(-\frac{y_1^2+y_2^2+y_3^2+y_4^2+a}{2\sigma^2}\right)I_0(\frac{a\sqrt{y_3^2+y_4^2}}{\sigma^2})$$

Since $I_0(z)$ is an increasing function of $|z|$, the ML is of the stated form.
Another way to be convince yourself is to compute the LLR. We know that given $H=i,(i=0,1)$, $Y\sim\mathcal{N}(0,\Sigma_i)$ where:

$$\Sigma_0 = \begin{pmatrix} a^2+\sigma^2 & 0 & 0 & 0 \\ 0 & a^2+\sigma^2 & 0 & 0 \\ 0 & 0 & \sigma^2 & 0 \\ 0 & 0 & 0 & \sigma^2 \end{pmatrix} \quad \Sigma_1 = \begin{pmatrix} \sigma^2 & 0 & 0 & 0 \\ 0 & \sigma^2 & 0 & 0 \\ 0 & 0 & a^2+\sigma^2 & 0 \\ 0 & 0 & 0 & a^2+\sigma^2 \end{pmatrix}$$

The pdf of Y given $H=0$ is given by:

$$f_{Y|H}(y|0) = \frac{1}{(2\pi)^2\sigma^2(a^2+\sigma^2)}e^{-\frac{1}{2}\left[\frac{y_1^2+y_2^2}{a^2+\sigma^2}+\frac{y_3^2+y_4^2}{\sigma^2}\right]}$$

We can also write $f_{Y|H}(y|1)$ and see that it depends only on v_0 and v_1.
Now we can write the LLR and find that it is

$$LLR(Y) = \frac{a^2}{2\sigma^2(a^2+\sigma^2)}(v_0-v_1)$$

We will thus choose $H = 0$ whenever $v_0 > v_1$ and $H = 1$ otherwise.
(b) Given $H = 0$, $V_0 = Y_1^2 + Y_2^2$, the distribution of which we found in Exercise 2.1:

$$f_{V_0|H}(v|0) = \frac{1}{2(a^2+\sigma^2)} \exp(-v/(2(a^+\sigma^2))), v \geq 0$$

Similarly $V_0 = Y_1^2 + Y_2^2$ has distribution.

$$f_{V_1|H}(v|0) = \frac{1}{2\sigma^2} \exp(-v/(2\sigma^2)), v \geq 0$$

(c) Let $U = V_0 - V_1$ compute $P(U > u|H = 0)$. Conditioning on V_1 we have

$$\begin{aligned}
P(U > u|H = 0) &= \int_0^\infty P(V_0 > u + v|H = 0, V_1 = v) f_{V_1|H}(v|0) \\
(V_0 \text{ ind. } V_1) &= \int_0^\infty P(V_0 > u + v|H = 0) f_{V_1|H}(v|0) \\
&= \begin{cases} \frac{a^2+\sigma^2}{2\sigma^2+a^2} \exp(-\frac{1}{2(a^2+\sigma^2)}u), & u \geq 0 \\ 1 - \frac{\sigma^2}{2\sigma^2+a^2} \exp(\frac{1}{2\sigma^2}u), & u < 0 \end{cases}
\end{aligned}$$

Thus the pdf is given by:

$$f_{U|H}(u|0) = \begin{cases} \frac{1}{2(2\sigma^2+a^2)} \exp(-\frac{1}{2(a^2+\sigma^2)}u), & u \geq 0 \\ \frac{1}{2(2\sigma^2+a^2)} \exp(\frac{1}{2\sigma^2}u), & u < 0 \end{cases}$$

(d) Given $H = 0$, an error occurs whenever $U < 0$. The corresponding probability is

$$\begin{aligned}
P(U < 0|H = 0) &= \int_{-\infty}^0 P_{U|H}(u|0) \\
&= \int_{-\infty}^0 \frac{1}{2(2\sigma^2+a^2)} \exp(\frac{1}{2\sigma^2}u) du \\
&= \frac{\sigma^2}{2\sigma^2+a^2} \\
&= (2 + a^2/\sigma^2)^{-1}
\end{aligned}$$

(e) By symmetry of the transmitted signals and the noise, this is also the overall error probability.