Lecture 4:

Sept. 28, 2010

<u>UMP</u> (Uniformly Most Powerful) Test

In Neyman-Pearson test, one tries to find <u>the most powerful</u> test, i.e., one that has the highest probability of detection ($P_D(\delta)$) for a given level, i.e., a given bound α on $P_{FA}(\delta)$ (probability of false alarm).

- In a composite hypothesis testing problem if, one can find a test that maximizes $P_D(\delta;\theta)$ for all $\theta \in \Lambda_1$ subject to $P_F(\delta;\theta) \leq \alpha$, $\theta \in \Lambda_0$ we say that there is a <u>u</u>niformly <u>m</u>ost <u>p</u>owerful (UMP) test.

The following example clarifies the idea.

<u>Example</u>: Testing the radius of a Point in a Plane.

Assume that a point is either at the origin or on a circle of radius A. Therefor, its coordinates are either $(0,0)$ or

$(A\cos\varphi, A\sin\varphi)$. Further assume that we observe the measurement ~~is~~ corrupted by a Gaussian error $(\varepsilon_1, \varepsilon_2) \sim \mathcal{N}\left(\{0,0\}, \begin{bmatrix}\sigma^2 & 0\\ 0 & \sigma^2\end{bmatrix}\right)$

So, the hypothesis are:

$$H_0: \begin{cases} Y_1 = \varepsilon_1 \\ Y_2 = \varepsilon_2 \end{cases}$$

versus

$$H_1: \begin{cases} Y_1 = A\cos\varphi + \varepsilon_1 \\ Y_2 = A\sin\varphi + \varepsilon_2 \end{cases}$$

In this problem $\underline{y} = (y_1, y_2) \in \mathbb{R}^2$, i.e., $\Gamma = \mathbb{R}_2$

The parameter $\underline{\theta} = (\theta_1, \theta_2)$ where

$\theta_1 \in \{0, A\}$, i.e., a constant being either equal to 0 or A.

$\theta_2 \in [0, 2\pi]$, Say uniformly distributed betwe 0 and 2π.

So $\Lambda = \{0, A\} \times [0, 2\pi]$,

and

$$\Lambda_0 = \{\underline{\theta} \in \Lambda \mid \theta_1 = 0\}$$

$$\Lambda_1 = \{\underline{\theta} \in \Lambda \mid \theta_1 = A\}$$

The conditional density $p(\underline{y}|\underline{\theta})$ is:

$$p(\underline{y}|\underline{\theta}) = \frac{1}{2\pi\sigma^2} \exp\left[-\frac{(y_1-\theta_1\cos\theta_2)^2+(y_2-\theta_1\sin\theta_2)^2}{2\sigma^2}\right]$$

For H_0,

$$p(\underline{y}|\underline{\theta}\in\Lambda_0) = p(\underline{y}|\underline{\theta})\Big|_{\theta_1=0}$$

$$= \frac{1}{2\pi\sigma^2}\exp\left[-\frac{(y_1^2+y_2^2)}{2\sigma^2}\right]$$

and for H_1,

$$p(\underline{y}|\underline{\theta}\in\Lambda_1) = \frac{1}{2\pi}\int_0^{2\pi} p(\underline{y}|\underline{\theta})\Big|_{\theta_1=A}\, d\theta_2$$

$$= \frac{1}{4\pi^2\sigma^2}\int_0^{2\pi}\exp\left[-\frac{(y_1-A\cos\theta_2)^2+(y_2-A\sin\theta_2)^2}{2\sigma^2}\right]d\theta$$

The likelihood ratio is:

$$L(\underline{y}) = \frac{p(\underline{y}|\Lambda_1)}{p(\underline{y}|\Lambda_0)} = \frac{e^{-A^2/2\sigma^2}}{2\pi}\int_0^{2\pi}\exp\left[\frac{A}{\sigma^2}(y_1\cos\theta_2+y_2\sin\theta_2)\right]d\theta_2$$

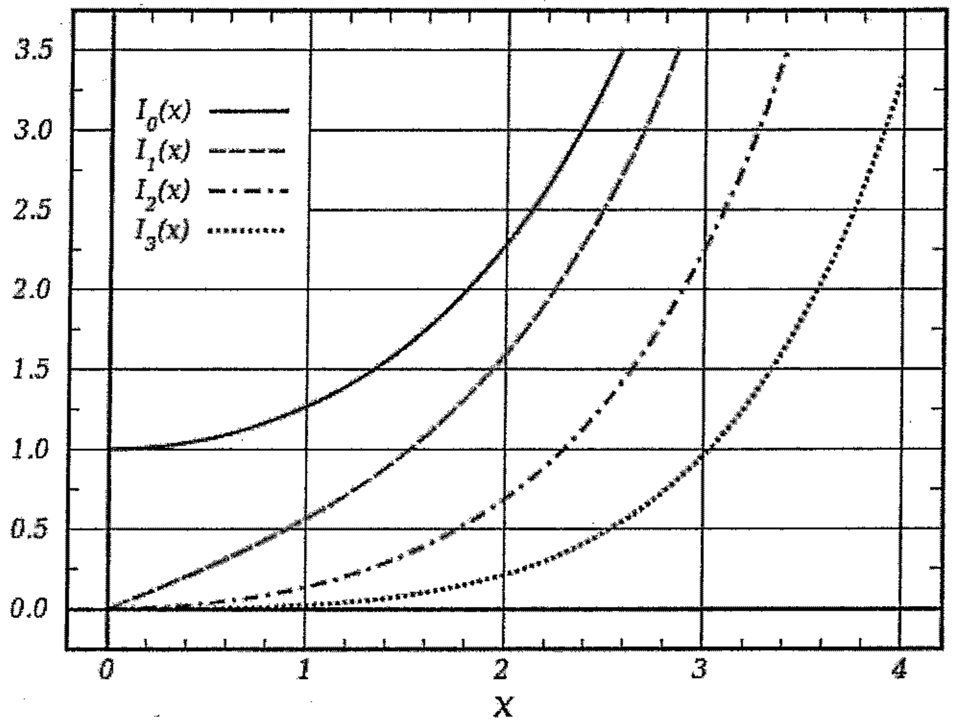
Let $r=\sqrt{y_1^2+y_2^2}$ and $\phi=\tan^{-1}\left(\frac{y_2}{y_1}\right)$. Then

$y_1 = r\cos\phi$ and $y_2 = r\sin\phi$ and,

$$L(y) = \frac{e^{-A^2/2\sigma^2}}{2\pi}\int_0^{2\pi}\left\{\frac{Ar}{\sigma^2}\cos(\theta_2-\phi)\right\}d\theta_2$$

$$= e^{-A^2/2\sigma^2} I_0\left(\frac{Ar}{\sigma^2}\right)$$

where $I_0(\cdot)$ is the modified Bessel function of zero-th order of the first kind.



we observe that since $I_0(x)$ is monotonically increasing with x, we can replace the $L(y)$, i.e., $I_0(\frac{Ar}{\sigma^2})$ with threshold, by a comparision of r with another threshold, i.e.,

$$\tilde{\delta}_0(y) = \begin{cases} 1 & \text{if } r > \tau' \\ \gamma & \text{if } r = \tau' \\ 0 & \text{if } \tau < \tau' \end{cases}$$

where $\tau' = \frac{\sigma^2}{A} I_0^{-1}(\tau e^{A^2/2\sigma^2})$.

The Decision Region for this problem is:

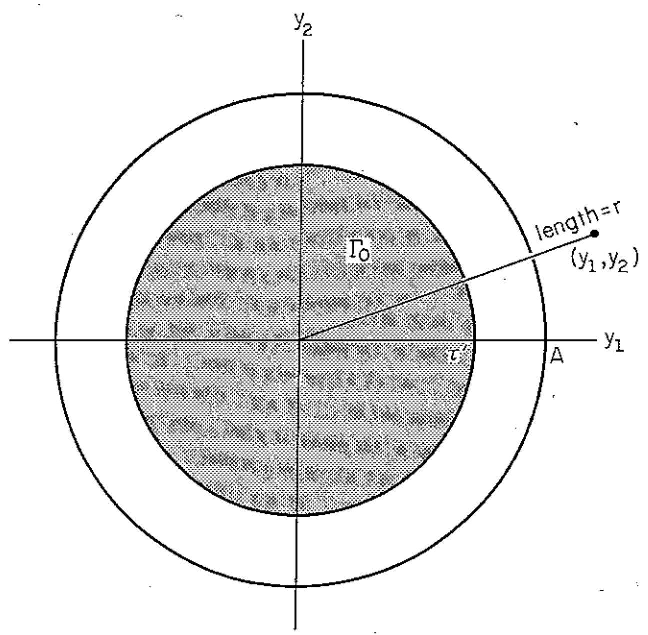


FIGURE II.E.1. Decision regions for Example II.E.1 ($\Gamma_1 = \Gamma_0^c$).

The above test does not depend on $\underline{\theta}$. So, it is a UMP test. UMP tests are desired since sometimes we do not have knowledge of the distribution of the paramaters.

For a randomized decision rule $\tilde{\delta}$, we can define

$$P_F(\tilde{\delta}, \underline{\theta}) = E_{\underline{\theta}}\{\tilde{\delta}(Y)\}, \quad \underline{\theta} \in \Lambda_0$$

and

$$P_D(\tilde{\delta}, \underline{\theta}) = E_{\underline{\theta}}\{\tilde{\delta}(Y)\}, \quad \underline{\theta} \in \Lambda_1$$

So, a UMP test of α-level exists if:

the optimum test is the one that maximizes $P_D(\tilde{\delta}; \underline{\theta})$ for every $\underline{\theta} \in \Lambda_1$ subject to the

Constraint:

$$\{ P_F(\tilde{\delta}, \underline{\theta}) \le \alpha, \ \theta \in \Lambda_0 \}.$$

While UMP is desirable, it ~~is~~ does not alway exist.

<u>Take the following example</u>:

Example: Testing of location with Gaussian erro

Consider

$$H_0 : \ \theta = \mu_0 \quad \Rightarrow Y = \mu_0 + \varepsilon$$

versus

$$H_1 : \ \theta > \mu_0 \quad \Rightarrow \ Y = \theta + \varepsilon \ , \ \theta > \mu_0$$

$$\varepsilon \sim N(0, \sigma^2)$$

Note that, here we have a simple null hypothesis $\Lambda_0 = \{\mu_0\}$ and a composite competing hypothesis $\Lambda_1 = (\mu_0, \infty)$, i.e. $\Lambda = [\mu_0, \infty)$. For each $\theta \in \Lambda$, we have $Y \sim N(\theta, \sigma^2)$ and

$$\Gamma_\theta = \{ y \in \Gamma \mid y > \sigma Q^{-1}(\alpha) + \mu_0 \}$$

and $P_D(\tilde{\delta}_1; \theta) = Q\left(Q^{-1}(\alpha) - \dfrac{\theta - \mu_0}{\sigma}\right)$

[Note: See lecture note 2 or page 36 of the text book].

So, this problem has a UMP test.

Now consider a problem that is a little bit different from the above, i.e.,

$$H_0: \quad \theta = \mu_0$$

versus

$$H_1: \quad \theta \neq \mu_0$$

In this problem,

$\Lambda_0 = \{\mu_0\}$ as before, but

$$\Lambda_1 = (-\infty, \mu_0) \cup (\mu_0, \infty)$$

For $\theta > \mu_0$, the decision rule (as we saw above) is:

$$\Gamma_\theta = \{ y \in \Gamma \mid y > \sigma Q^{-1}(Q(\alpha)) + \mu_0 \}$$

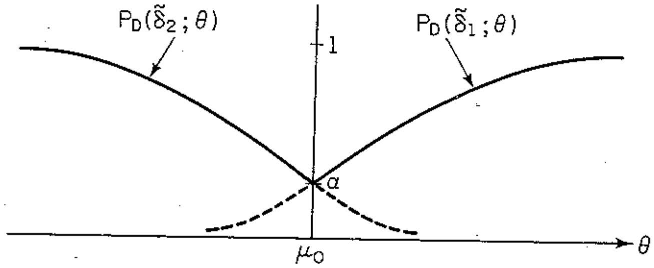
For $\theta < 0$, however,

$$\Gamma_\theta = \{ y \in \Gamma \mid y < Q^{-1}(1-\alpha) + \mu_0 \}$$

If we denote this decision rule by $\tilde{\delta}_2$,

$$P_D(\tilde{\delta}_2; \theta) = 1 - Q[Q^{-1}(1-\alpha) - \tfrac{\theta - \mu_0}{\sigma}]$$

Note that while we have UMP tests for $\theta > 0$ and $\theta < 0$, the two decision regions are different. So, there is no UMP for the whole range of θ, i.e., this problem has no UMP test



The above figure shows the power curves for test of $\theta = \mu_0$ versus $\theta > \mu_0$ and $\theta = \mu_0$ versus $\theta < \mu_0$.

In the cases like the above problem UMP criterion is too strong.

For some situations a locally most powerful (LMP) test is suitable.

For example, in cases where $\Lambda = [\theta_0, \infty)$ with $\Lambda_0 = \{\theta_0\}$ and $\Lambda_1 = (\theta_0, \infty)$, i.e.,

$$H_0 : \theta = \theta_0$$

versus

$$H_1 : \theta > \theta_0$$

This type of problem arises when we look to find whether the signal does not exist, i.e., $\theta_0 = 0$ and $\theta > \theta_0 = 0$ is the amplitude of the signal if it is present.

[Note: in the case of Gaussian error, this problem has a UMP test as we saw before, but, not in general.]

Most often in problems we are interested in cases where under H_1, θ is near θ_0, i.e., the cases where target is present but the amplitude is so low that is near absent.

Consider a decision rule $\tilde{\delta}$ and denote its power (probability of detection) for θ as $P_D(\tilde{\delta}; \theta)$. We can expand $P_D(\tilde{\delta}, \theta)$ using Taylor series as:

$$P_D(\tilde{\delta};\theta) = P_D(\tilde{\delta};\theta_0) + (\theta-\theta_0)P_D'(\tilde{\delta};\theta_0) + O((\theta-\theta_0)^2)$$

where

$$P_D'(\tilde{\delta};\theta) = \frac{\partial P_D(\tilde{\delta};\theta)}{\partial \theta}$$

and $O((\theta-\theta_0)^2)$ represents second order and higher order terms.

Note that

$$P_D(\tilde{\delta}; \theta_0) = P_F(\tilde{\delta})$$

That is probability of detection when there is no target is the probability of <u>false</u> <u>alarm</u>.

So for all size-α (or level-α) tests

$$P_D(\tilde{\delta}; \theta) \cong \alpha + (\theta - \theta_0) P_D'(\tilde{\delta}; \theta_0)$$

So, for θ near θ_0, we can maximize the power of an α-level test by maximizing

$$P_D'(\tilde{\delta}; \theta_0) = \left. \frac{\partial P_D(\delta, \theta)}{\partial \theta} \right|_{\theta = \theta_0}$$

This is called an α-level <u>locally</u> <u>most</u> <u>powerfull</u> (LMP) test.

Now let's see what the structure of an LMP test is. Note that

$$P_D(\tilde{\delta}; \theta) = E_\theta\{\tilde{\delta}(Y)\}$$

$$= \int_\Gamma \tilde{\delta}(y) p(y|\theta) dy$$

Then,

$$P_D'(\tilde{\delta};\theta_0) = \int_\Gamma \tilde{\delta}(y)\,\frac{\partial}{\partial\theta}p(y|\theta)\Big|_{\theta=\theta_0} dy$$

Comparing this with Neyman-Pearson test's

$$P_D(\tilde{\delta}) = \int_\Gamma \tilde{\delta}(y)\,p(y|H_1)\,dy$$

we see that finding α-level LMP test is similar to NP test if we replace $p(y|H_1)$ by $\frac{\partial}{\partial\theta}p(y|\theta)\Big|_{\theta=\theta_0}$, i.e.,

$$\tilde{\delta}_{lo}(y) = \begin{cases} 1 & \\ \gamma & \text{if } \frac{\partial}{\partial\theta}p(y|\theta)\Big|_{\theta=\theta_0} \begin{matrix} > \\ = \\ < \end{matrix} \eta\, p(y|\theta_0) \\ 0 & \end{cases}$$

- Example: Locally Optimum Detection of Coherent Signals in i.i.d. noise

In certain detection problems, we may know the form of the received signal but not its amplitude. This problem can be modelled as a composite hypothesis test as follows.

$$H_0: \; Y_k = N_k \quad k=1,2,\dots,n$$

versus

$$H_1: \; Y_k = N_k + \theta S_k \quad k=1,2,\dots,n \quad \theta>0$$

where $\underline{S}=(s_1,\dots,s_n)^T$ is a known signal

$\underline{N}=(N_1,\dots,N_n)^T$ is a continuous random noise with independent identically distributed (i.i.d.) distribution with marginal density $p_{N_k}(n_k)$ and θ is a parameter indicating the signal strength.

Given θ, the likelihood ratio of H_1 versus H_0 is

$$L_\theta(\underline{y}) = \prod_{k=1}^{n} \frac{p_{N_k}(y_k-\theta s_k)}{p_{N_k}(y_k)}$$

The critical region

$$\Gamma_\theta = \{\underline{y}\in \mathbb{R}^n \mid L_\theta(\underline{y}) > \tau\}$$

in general depends on θ. Except for some special cases (Gaussian noise case considered in last lecture is one special case) there is no UMP test for problems like this.

Let's consider the LMP test.

The locally optimum test for H_0 versus H_1 is:

$$\tilde{\delta}_{lo}(\underline{y}) = \begin{cases} 1 & \\ \gamma & \text{if } \frac{\partial}{\partial\theta} L_\theta(\underline{y})\Big|_{\theta=0} \gtreqless \tau \\ 0 & \end{cases}$$

where

$$\frac{\partial}{\partial\theta} L_\theta(\underline{y}) = \sum_{k=1}^{n} \frac{\partial}{\partial\theta} \frac{p_{N_k}(y_k-\theta s_k)}{p_{N_k}(y_k)}$$

$$= \sum_{k=1}^{n} \frac{1}{p_{N_k}(y_k)} \frac{\partial}{\partial\theta} p_{N_k}(y_k-\theta s_k)$$

$$= \sum_{k=1}^{n} \frac{1}{p_{N_k}(y_k)} \frac{\partial}{\partial\theta}(y_k-\theta s_k) \frac{\partial p_{N_k}(y_k)}{\partial y_k}$$

$$= \sum -s_k \frac{1}{p_{N_k}(y_k)} \frac{\partial p_{N_k}(x)}{\partial x}\Big|_{x=y_k}$$

$$= \sum_{k=1}^{n} s_k \, g_{lo}(y_k)$$

where

$$g_{lo}(x) = -\frac{\frac{d p_N(x)}{dx}}{p_N(x)}$$

The above can be applied to different distribution

A) In the Case of Gaussian noise

$$P_N(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}$$

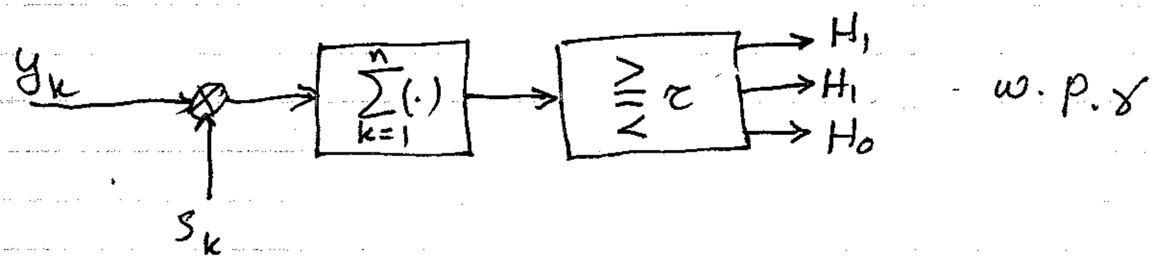
$$\frac{d}{dx} P_N(x) = -\frac{x}{\sigma^2}\left(\frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2}\right) = \frac{-x}{\sigma^3\sqrt{2\pi}} e^{-x^2/2\sigma^2}$$

$$g_{lo}(x) = -\frac{d/dx\, g_N(x)}{g_N(x)} = \frac{x}{\sigma^2}$$

So,

$$\frac{\partial}{\partial\theta} L_\theta(\underline{y}) = \frac{1}{\sigma^2}\sum_{k=1}^{n} s_k y_k$$

This results in the correlation receiver



B) In the case of Laplacian Noise

$$P_N(x) = \frac{\alpha}{2} e^{-\alpha|x|} \qquad x \in \mathbb{R}$$

for $x < 0 \Rightarrow P_N(x) = \frac{\alpha}{2} e^{\alpha x} \Rightarrow \frac{\partial P_N(x)}{\partial x} = \frac{\alpha^2}{2} e^{\alpha x}$

and

$$g_{lo}(x) = -\frac{\alpha^2/2\, e^{\alpha x}}{\frac{\alpha}{2} e^{\alpha x}} = -\alpha$$

For $\alpha \geqslant 0$

$$p_N(x) = \frac{\alpha}{2} e^{-\alpha x} \Rightarrow \frac{\partial p_N(x)}{\partial x} = -\frac{\alpha^2}{2} e^{-\alpha x}$$

$$g_{l0}(x) = +\alpha$$

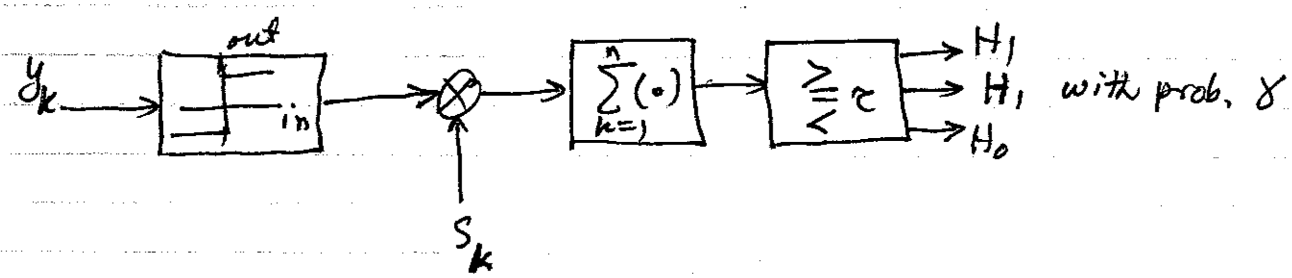
So,

$$g_{l0}(x) = \alpha \, \mathrm{Sgn}(x)$$

and

$$\frac{\partial}{\partial} L_\theta(\underline{y}) = \alpha \sum_{k=1}^{n} s_k \, \mathrm{Sgn}(y_k)$$

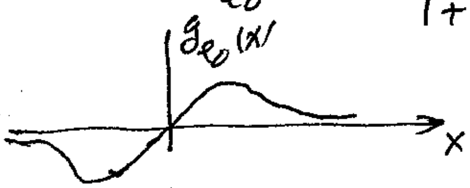
So, we need to pass $\{y_k\}$ through a limiter and then multiply by $\{s_k\}$ and accumulate:



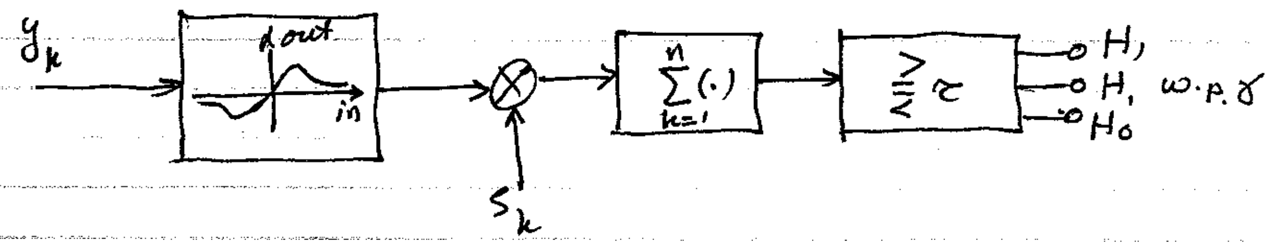
c) Cauchy distributed noise:

$$p_N(x) = \frac{1}{\pi(1+x^2)} \qquad \forall x \in \mathbb{R}$$

In this case $\quad g_{l0}(x) = \dfrac{2x}{1+x^2}$



In this case, the function $g_{lo}(x)$ is linear near zero and zero for large values of x. This means that a locally optimal detector ignores the extreme values



$g_{lo}(x)$ can be approximated by:

$$g_{lo}(x) = \begin{cases} x & \text{if } |x| \leq K \\ 0 & \text{if } |x| > K \end{cases}$$