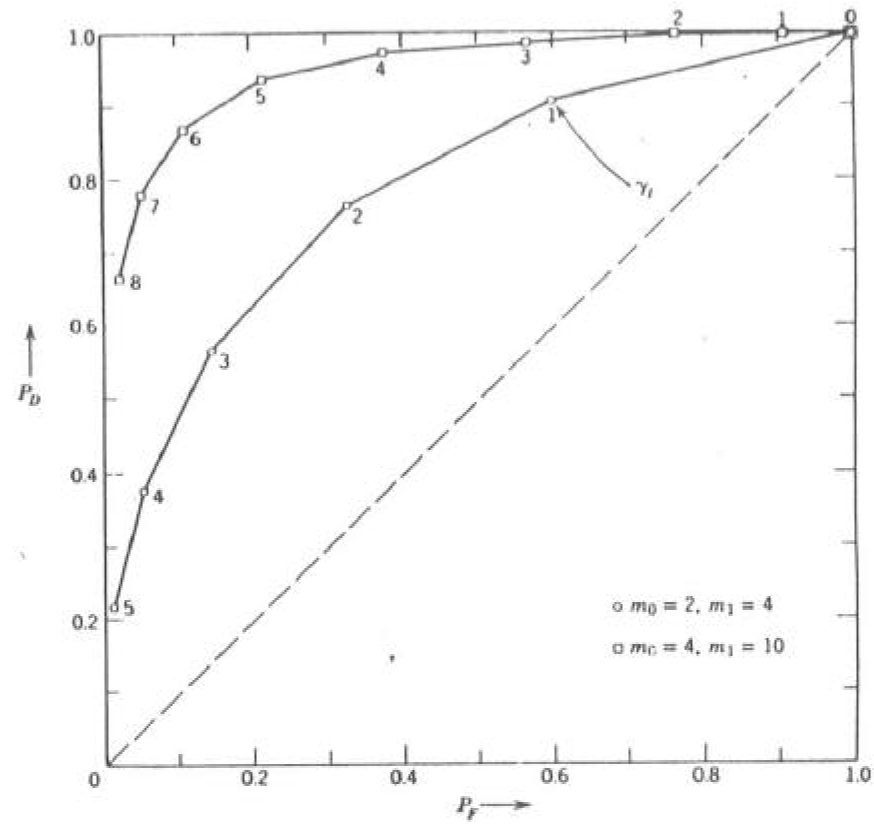
ROC for randomized tests



* Consists of straight lines which connect discrete points

In general

$$P_F(\eta) = \int_h^\infty p(\Lambda|H_0)\, d\Lambda$$

If $P_F(\eta)$ is a continuous function of η, we can achieve a desired value from 0 to 1 by a suitable choice of η and a randomized test will never be needed.

Properties of ROC

1) All continuous LRT's have ROC's that are concave downward. If they were not, a randomized test would be better. This would contradict proof that a LRT is optimum.
2) All the LRT's have ROC's that are above the $P_D = P_F$ line. Special case of 1 as ($P_F = 0, P_D = 0$) and ($P_F = 1, P_D = 1$) are in all ROC's.
3) The slope of a curve in a ROC at a particular point is equal to the value of the threshold η required to achieve the P_D and P_F of that point.

Proof

$$P_D = \int_{h}^{\infty} p(\Lambda|H_1)\, d\Lambda$$

$$P_F = \int_{h}^{\infty} p(\Lambda|H_0)\, d\Lambda$$

$$\Rightarrow \frac{dP_D / dh}{dP_F / dh} = \frac{-p(h \mid H_1)}{-p(h \mid H_0)} = \frac{dP_D}{dP_F}$$

To show that $\frac{p(h \mid H_1)}{p(h \mid H_0)} = \eta$

Let $\Omega(\eta) = \{ \mathbf{R} | \Lambda(\mathbf{R}) \geq \eta \} = [\mathbf{R} | \frac{p(h \mid H_1)}{p(h \mid H_0)} \geq \eta]$

Then

$$P_D(\eta) = \Pr\{\Lambda(\mathbf{R}) \geq \eta | H_1\}$$

$$= \int_{\Omega(h)} p(\mathbf{R}|H_1)\, d\mathbf{R}$$

$$= \int_{\Omega(h)} \Lambda(\mathbf{R})\, p(\mathbf{R}|H_1)\, d\mathbf{R}$$

Using the definition of $\Omega(\eta) = \{ \mathbf{R} | \Lambda(\mathbf{R}) \geq \eta \}$,

$$P_D(\eta) = \int_{h}^{\infty} \Lambda\, p(\Lambda|H_0)\, d\Lambda$$

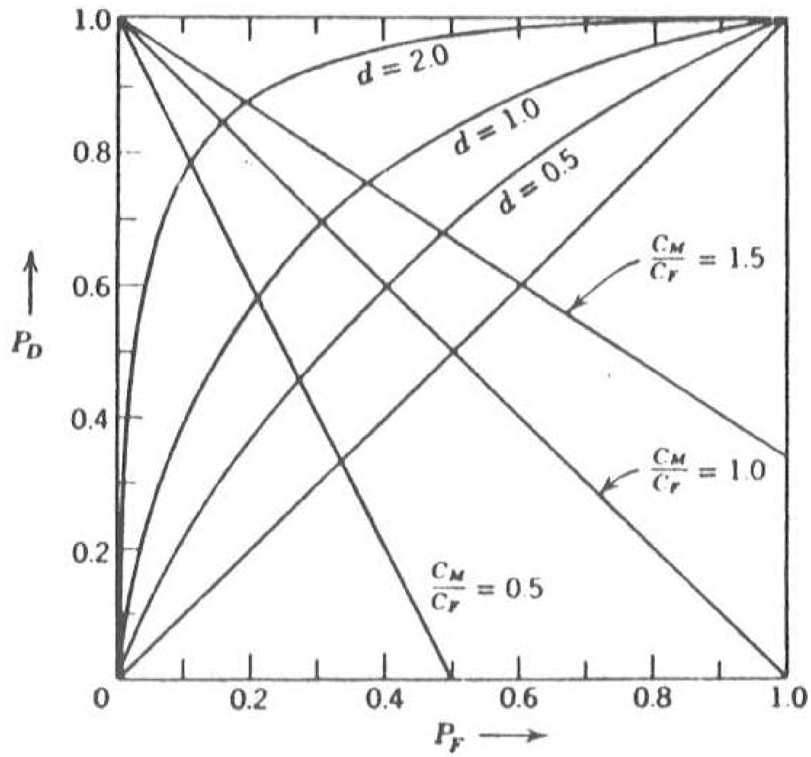
$$\Rightarrow \quad \frac{dP_D(h)}{dh} = -\eta\, p(\eta|H_0) \qquad \text{QED}$$

Property 4 : Whenever the maximum value of the Bayes Risk is interior to the interval (0, 1) on the P_1 axis, the minimax operating point is the intersection of the line.

$$(C_{11}-C_{00}) + (C_{01}-C_{11})(1-P_D) - (C_{10}-C_{00}) P_F = 0$$

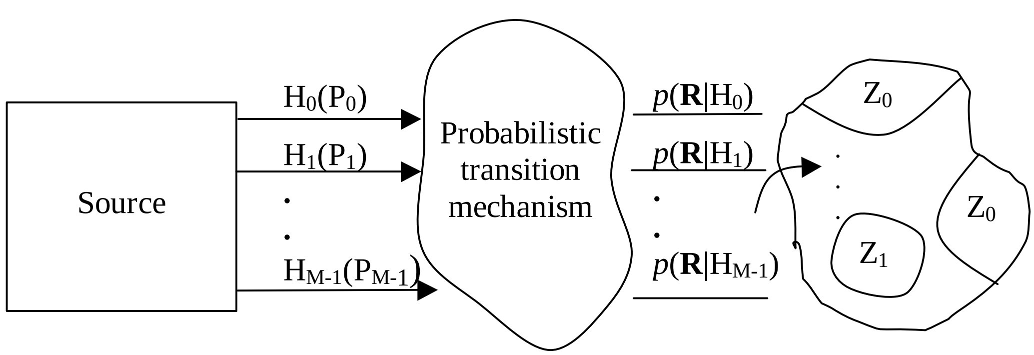
and the appropriate curve of ROC.

Special case $C_F P_F = C_M P_M = C_M (1 - P_D)$



- Using either a Bayes test or a Neyman-Pearson criterion, optimum test is a LRT. Then, regardless of the dimensionality of the observation space, the test consists of comparing a seatar variable $\Lambda(\mathbf{R})$ to a threshold. (P_F/η) cts.)

M Hypothesis
M-ary Bayes Receiver



Costs – C_{ij} – ncost of choosing H_i when H_j is true.

$$\mathsf{R} = \sum_{j=0}^{M-1}\sum_{i=0}^{M-1} P_j C_{ij} \Pr[\text{choosing } H_i | H_j \text{ is true}]$$

$$= \sum_{j=0}^{M-1}\sum_{i=0}^{M-1} P_j C_{ij} \int_{Z_i} p(\mathbf{R}|H_j)\, d\mathbf{R}$$

- partition Z so that R is minimized.

Consider $C_{ij} = 0, \quad i=j$
$C_{ij} = C, \quad i \neq j$

These indicate that any error is of equal importance.
$\Rightarrow$ minimizing the total probability of error.

Aside $p(x|A) = \dfrac{\Pr[A|x]\, p_x(x)}{\Pr[A]}$

$\therefore$ The Bayes risk can be written as :

$$\mathsf{R} = \sum_{j=0}^{M-1}\sum_{i=0}^{M-1} P_j C_{ij} \int_{Z_i} \frac{p(\mathbf{R})\Pr[H_j|\mathbf{R}]}{P_j}\, d\mathbf{R}$$

$$= C \sum_{j=0}^{M-1} \int_{Z_i} \underbrace{\left(\sum_{\substack{j=0 \\ j\neq i}}^{M-1} \Pr[H_j | \mathbf{R}] \right)}_{1 - \Pr[H_i|\mathbf{R}]} p(\mathbf{R}) d\mathbf{R}$$

$$\mathsf{R} = C \left(\underbrace{\sum_{j=0}^{M-1} \int_{Z_i} p(\mathbf{R}) d\mathbf{R}}_{=1} - \left(\int_{Z_0} \Pr[H_0 | \mathbf{R}]\, p(\mathbf{R}) d\mathbf{R} + \ldots + \int_{Z_{M-1}} \Pr[H_{M-1} | \mathbf{R}]\, p(\mathbf{R}) d\mathbf{R} \right) \right)$$

Since $\Pr[H_i|\mathbf{R}] \geq 0, \quad \forall\, \mathbf{R}$
$P(\mathbf{R}) \geq 0, \quad \forall\, \mathbf{R}$

To minimize R assign **R** to region for which $\Pr[H_i|\mathbf{R}]$ is maximum.
$\Rightarrow$ maximum a posteriori MAP $\rightarrow$ receiver

If M hypothesis are equally likely

$P_j = \dfrac{1}{M} \quad \forall$, then

$$\Pr[H_j|\mathbf{R}] = \frac{p(\mathbf{R}\,|\,H_j)\Pr[H_j]}{p(\mathbf{R})} \quad \Rightarrow \text{independent of j}$$

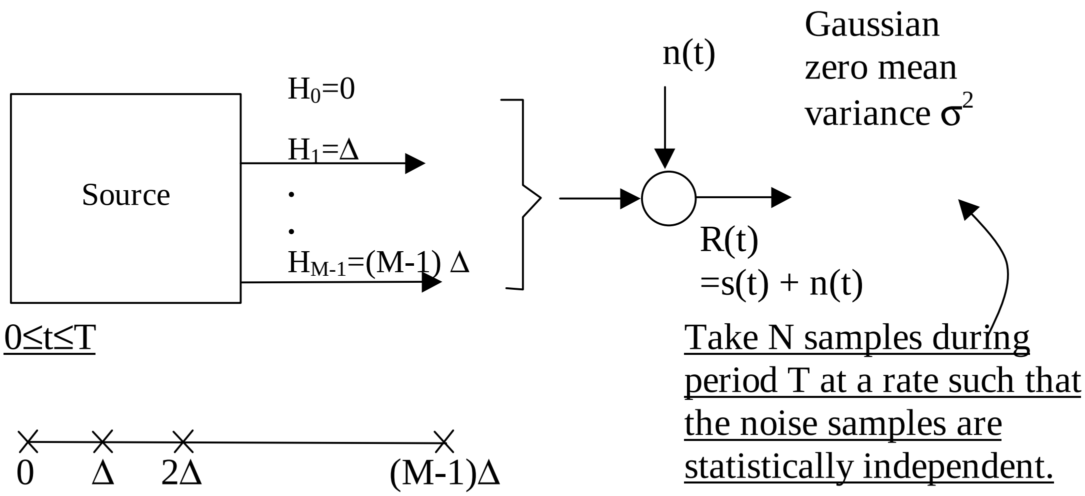
(arrow pointing to $\Pr[H_j]$: $\frac{1}{M}$)

∴ Compute $p(\mathbf{R}|H_j)$ and choose hypothesis for which this is maximum.
$j=0, 1, \ldots, M-1$

⇒ Maximum likelihood-ML-Receiver.

<u>Example</u>

<u>**M-ary ASK**</u>



$\Pr[H_i] = \frac{1}{M}$ $\forall i$, $C_{ii} = 0$, $C_{ij} = 1$, $i \neq j$

$R_1 = R(t_1) = i\Delta + n(t_1)$ $\quad R_N = R(t_N) = i\Delta + n(t_N)$

$R_2 = R(t_2) = i\Delta + n(t_2)$

.

.

$R_j = R(t_j) = i\Delta + n(t_j)$ $\quad i = 0, 1, 2, \ldots, M-1$

Receiver – Compute

$\Pr[H_i|R_1, R_2, \ldots, R_N]$ or $\frac{p(R_1, R_1, \ldots, R_N\,|\,H_i)\Pr[H_i]}{p(\mathbf{R})}$

or $\quad p(R_1, R_2, \ldots, R_N|H_i)$ <u>for i = 0, 1, 2, ..., M-1 and choose largest</u>

$$p(R_1, R_2, ..., R_N|H_i) = \prod_{j=1}^{N} \frac{1}{\sqrt{2ps}} e^{-(R_j - i\Delta)^2}$$

<u>Take ln</u>

$$= \underbrace{ln\left\{\left(\frac{1}{\sqrt{2ps}}\right)^N\right\}}_{\text{ignore}} - \frac{1}{2s^2} \sum_{j=1}^{N} (R_j - i\Delta)^2$$

$\therefore$ Compute $\sum_{j=1}^{N} (R_j - i\Delta)^2$ for i = 0, 1, 2, ..., M-1 and <u>choose smallest</u>

or compute $\sum_{j=1}^{N} 2R_j(i\Delta) - N(i\Delta)^2$ for i = 0, 1, 2, ..., M-1 and

<u>choose largest</u>

or $i\left[\frac{1}{N}\sum_{j=1}^{N} R_j - \frac{i\Delta}{2}\right]$, $\forall$ i choose largest.

Hypothesis H_i is accepted if when tested against H_{i-l} (l = 1, 2, ..., i-1) and against H_{i+l} (l = 1, 2, ..., M-1-i) the expression is largest, i.e.,

<u>H_i versus H_{i-l}</u> Reject H_{i-l}

$$i\left[\frac{1}{N}\sum_{j=1}^{N} R_j - \frac{i\Delta}{2}\right] > (i-l)\left[\frac{1}{N}\sum_{j=1}^{N} R_j - \frac{(i-l)\Delta}{2}\right]$$

$$\Rightarrow \quad \text{if } \frac{1}{N}\sum_{j=1}^{N} R_j > \left(i - \frac{l}{2}\right)\Delta$$

$$\Rightarrow \quad \frac{1}{N}\sum_{j=1}^{N} R_j > \left(i - \frac{1}{2}\right)\Delta$$ reject H_{i-l} ; l = 1, 2, ..., i-1, i

Similarly for H_i versus H_{i+l} we get

$$\text{if } \frac{1}{N}\sum_{j=1}^{N} R_j < \left(i + \frac{l}{2}\right)\Delta$$

$$\Rightarrow \quad \frac{1}{N}\sum_{j=1}^{N} R_j < \left(i + \frac{1}{2}\right)\Delta$$ reject H_{i+l} ; l = 1, 2, ..., M-1-i

$\therefore$ If $\left(i - \frac{1}{2}\right)\Delta < \frac{1}{N}\sum_{j=1}^{N} R_j < \left(i + \frac{1}{2}\right)\Delta$, i = 1, 2, ..., M-2 <u>accept H_i</u>

i = 0, then accept H_0 if $\frac{1}{N}\sum_{j=1}^{N} R_j < \frac{\Delta}{2}$,

i = M-1, then accept H_{M-1} if $\frac{1}{N}\sum_{j=1}^{N} R_j > \left(M-1-\frac{1}{2}\right)\Delta$

$$> \left(M-\frac{3}{2}\right)\Delta$$

$\frac{1}{N}\sum_{j=1}^{N} R_j$ is sufficient.

Statistic ⇒ knowledge of only this function of the observables allows one to distinguish among hypotheses.

Receiver Performance

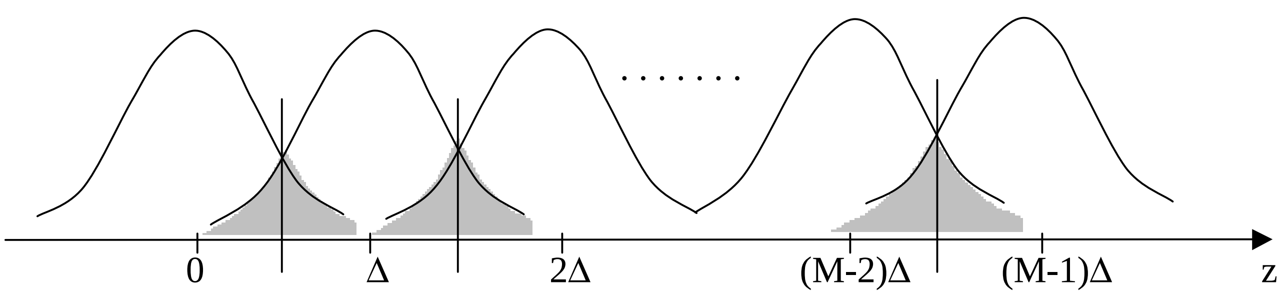
$$z = \frac{1}{N}\sum_{j=1}^{N} R_j$$

$$\Pr[\text{error}] = \sum_{i=0}^{M-1}\Pr[\text{error} \mid H_i]\Pr[H_i] = \frac{1}{M}\sum_{i=0}^{M-1}\Pr[\text{error} \mid H_i]$$

$$\Pr[\text{error}|H_i] = 1 - \Pr[\text{correct}|H_i] = 1 - \int_{\left(i-\frac{1}{2}\right)\Delta}^{\left(i+\frac{1}{2}\right)\Delta} p[z \mid H_i]\,dz$$

$$z - \text{Gaussian } E[z|H_i] = E\left[\frac{1}{N}\sum_{j=1}^{N}\left(i\Delta + n_j\right)\right] = i\Delta$$

$$\text{var}[z|H_i] = \frac{1}{N^2}\,.\,\sigma^2.\,N = \frac{s^2}{N} = \sigma_z^2$$



2(M-2) +2 = 2(M-1) areas

$$\therefore \quad \Pr[\text{error}] = \frac{2(M-1)}{M} \int_{\Delta/2}^{\infty} \frac{1}{\sqrt{2p}\,s_z} e^{-\frac{z^2}{2s_z^2}} dz$$

$$= \frac{2(M-1)}{M} Q\left(\frac{\Delta}{2s_z}\right)$$

$$= \frac{2(M-1)}{M} Q\left(\frac{\Delta\sqrt{N}}{2s}\right)$$

$\frac{\Delta\sqrt{N}}{s}$ can be interpreted as a signal to noise ratio.

Estimation Theory

- parameter estimation
- parameter space. The output of the source is a parameter (or variable).

Example $\quad r = a + n$

$\qquad\qquad -\infty < A < \infty$ or $-v \le A \le v$

The problem is to observe r and estimate a.

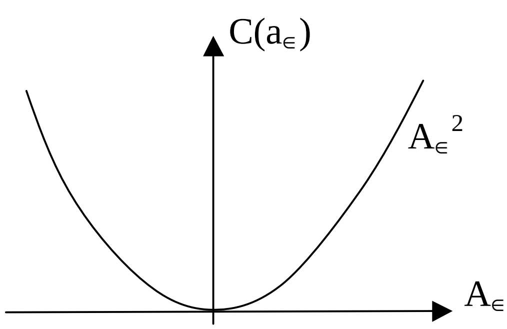
After observing $\mathbf{R} \Rightarrow$ estimate $a = \hat{a}(\mathbf{R})$.

Cases : 1) The parameter is a random variable – probability density
2) parameter unknown but not a random variable.

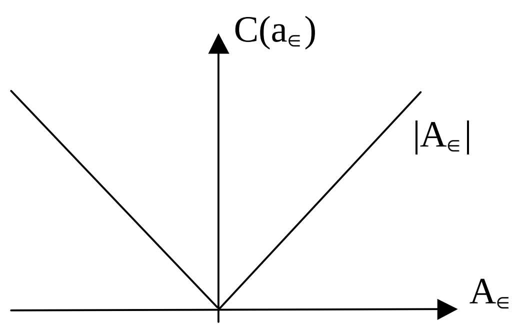
1) Random Parameters : Bayes Estimation

a, $\hat{a}(\mathbf{R})$ continuous variables.
Assign a cost to all pairs $[a, \hat{a}(\mathbf{R}).]$.

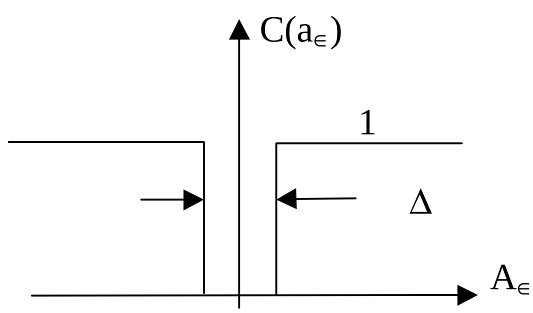
$\Rightarrow \quad a_{\in}(\mathbf{R}) \cong \hat{a}(\mathbf{R}) - a \Rightarrow C(a_{\in})$
function of one variable



Mean – square error



Absolute error



Typical cost function

Uniform cost function

Assume $p_a(A)$ is known.

$$\mathcal{R} \cong E\{C[a,\hat{a}(\mathbf{R})]\} = \int_{-\infty}^{\infty} dA \int_{-\infty}^{\infty} C[A,\hat{a}(\mathbf{R})]\, p(A,\mathbf{R})\, d\mathbf{R}$$

$$\Rightarrow \qquad \mathcal{R} \cong \int_{-\infty}^{\infty} dA \int_{-\infty}^{\infty} C[A-\hat{a}(\mathbf{R})]\, p_{a,\mathbf{r}}(A,\mathbf{R})\, d\mathbf{R}$$

* <u>mean – square</u>

$$\mathcal{R}_{ms} \cong \int_{-\infty}^{\infty} dA \int_{-\infty}^{\infty} d\mathbf{R}[A-\hat{a}(\mathbf{R})]^2\, p_{a,\mathbf{r}}(A,\mathbf{R}) \qquad (**)$$

Now,

$$p_{a,\mathbf{r}}(A, \mathbf{R}) = p_{\mathbf{r}}(\mathbf{R}) \,.\, p(A|\mathbf{R})$$

Thus,

$$\mathcal{R}_{ms} \cong \int_{-\infty}^{\infty} dA\, \underbrace{p_{\mathbf{r}}(\mathbf{R})}_{\text{non(-)ve}} \int_{-\infty}^{\infty} \underbrace{dA[A-\hat{a}(\mathbf{R})]^2\, p(A|\mathbf{R})}_{\text{non(-)ve}}$$

$\Rightarrow$ minimize $\mathcal{R}_{ms}$ by minimizing the inner integral

$$\frac{d}{d\hat{a}} \int_{-\infty}^{\infty} dA[A-\hat{a}(\mathbf{R})]^2\, p(A|\mathbf{R})$$

$$= -2\int_{-\infty}^{\infty} A\, p(A|\mathbf{R})dA + 2\hat{a}(\mathbf{R}) \underbrace{\int_{-\infty}^{\infty} p(A|\mathbf{R})dA}_{=1}$$

$\equiv 0$

$$\Rightarrow \qquad \hat{a}_{ms}(\mathbf{R}) = \underbrace{\int_{-\infty}^{\infty} A\, p(A|\mathbf{R})dA}$$

Since $\frac{d^2}{d\hat{a}^2}(.) = 2 > 0$ This is a unique minimum.

mean of the a posteriori density
- conditional mean

From (**) the inner integral is the conditional variance for $\hat{a}_{ms}(\mathbf{R})$. Therefore, the minimum value of $\mathcal{R}_{ms}$ is the just the average of the conditional variance over all observations **R**.

* <u>absolute value</u>

$$\mathsf{R}_{abs} = \int_{-\infty}^{\infty} d\mathbf{R}\, p(\mathbf{R}) \int_{-\infty}^{\infty} dA \left[\left| A - \hat{a}(\mathbf{R}) \right| \right] p(A \mid \mathbf{R})$$

To minimize the inner integral

$$I(\mathbf{R}) = \int_{-\infty}^{\hat{a}(\mathbf{R})} dA \left[\hat{a}(\mathbf{R}) - A \right] p(A \mid \mathbf{R}) + \int_{\hat{a}(\mathbf{R})}^{\infty} dA \left[A - \hat{a}(\mathbf{R}) \right] p(A \mid \mathbf{R})$$

$$\frac{d^2}{d\hat{a}^2}(.) = 0$$

$$\Rightarrow \int_{-\infty}^{\hat{a}_{abs}(\mathbf{R})} dA\, p(A \mid \mathbf{R}) = \int_{\hat{a}_{abs}(\mathbf{R})}^{\infty} dA\, p(A \mid \mathbf{R})$$

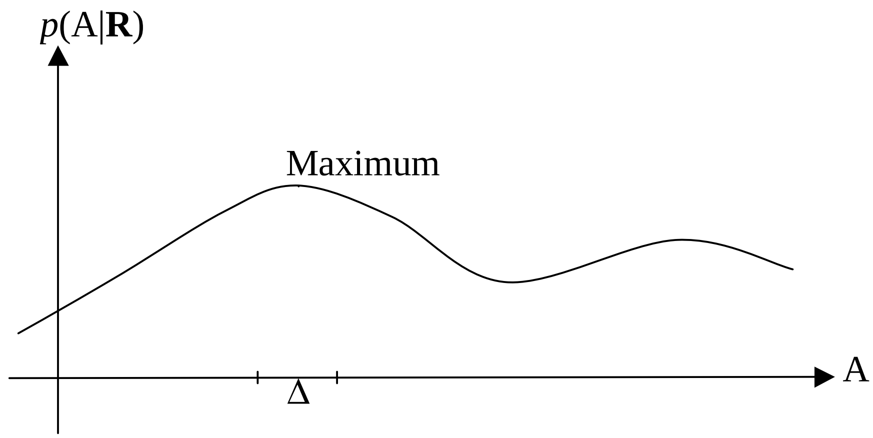
This is the definition of the "<u>median</u>" of the a posteriori density.

* <u>uniform cost</u>

$$\mathsf{R}_{unf} = \int_{-\infty}^{\infty} d\mathbf{R}\, p(\mathbf{R}) \left[1 - \int_{\hat{a}_{unf}(\mathbf{R}) - \Delta/2}^{\hat{a}_{unf}(\mathbf{R}) + \Delta/2} p(A \mid \mathbf{R}) dA \right]$$

To minimize this equation we maximize the inner integral.

Consider Δ arbitrarily small but nonzero.



⇒ for small Δ the best choice for $\hat{a}(\mathbf{R})$ is the value of A at which the a posteriori density has its maximum.

⇒ $\hat{a}_{map}(\mathbf{R})$: the "maximum a posteriori" estimate.

Need location of the maximum of $p(A|\mathbf{R})$.
Thus a necessary but not a sufficient condition.

$$\Rightarrow \quad \frac{\partial}{\partial A} \ln p(A \mid \mathbf{R}) \Big|_{A = \hat{a}(\mathbf{R})} = 0 \text{ – MAP Equation}$$

* In each case we must check to see if the solution is the absolute maximum.

Now, $p(\mathrm{A}|\mathbf{R}) = \dfrac{p(\mathbf{R}\,|\,\mathrm{A})\, p_a(\mathrm{A})}{\Pr(\mathbf{R})}$

$$\ln p(\mathrm{A}|\mathbf{R}) = \ln p(\mathbf{R}|\mathrm{A}) + \ln p_a(\mathrm{A}) - \ln \Pr(\mathbf{R})$$

ln Pr(**R**) is not a function of A.

$\Rightarrow l(\mathrm{A}) = \ln p(\mathbf{R}|\mathrm{A}) + \ln p_a(\mathrm{A})$

- $\ln p(\mathbf{R}|\mathrm{A})$ → probabilistic dependence of **R** on A
- $\ln p_a(\mathrm{A})$ → a priori knowledge

Thus MAP Equation

$$\frac{\partial l(\mathrm{A})}{\partial \mathrm{A}}\bigg|_{\mathrm{A}=\hat{a}(\mathbf{R})} = \frac{\partial \ln p(\mathbf{R}\,|\,\mathrm{A})}{\partial \mathrm{A}}\bigg|_{\mathrm{A}=\hat{a}(\mathbf{R})} + \frac{\partial \ln p_a(\mathrm{A})}{\partial \mathrm{A}}\bigg|_{\mathrm{A}=\hat{a}(\mathbf{R})} = 0$$

- minimum mean square error (MMSE)
- maximum a posteriori (MAP)

estimates

1. The minimum mean-square error estimate (MMSE) is always the mean of the a posteriori density (the conditional mean.)
2. MAP estimate is the value of A at which the a posteriori density has its maximum.
3. Optimum estimate is the conditional mean whenever the a posteriori density is a unimodal functional mean.

Many cases MMSE, MAP estimates equal.

Example

$$r_i = a + n_i\ ,\ i = 1, 2, ..., N$$

- a → $N(0, \sigma_a)$
- n_i → each independent $N(0, \sigma_n)$

$$p(\mathbf{R}|\mathrm{A}) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi}\,\sigma_n} exp\left[-\frac{(\mathrm{R}_i - \mathrm{A})^2}{2\sigma_n^2}\right]$$

$$p_a(\mathrm{A}) = \frac{1}{\sqrt{2\pi}\,\sigma_a} exp\left(-\frac{\mathrm{A}^2}{2\sigma_a^2}\right)$$

$$\hat{a}_{ms}(\mathbf{R}) \Rightarrow p(A|\mathbf{R})\ ? \overset{\text{know}}{\Rightarrow} \int_{-\infty}^{\infty} p(A|\mathbf{R})dA = 1$$

Thus,

$$p(A|\mathbf{R}) = \frac{p(\mathbf{R}\,|\,A)\, p_a(A)}{p(\mathbf{R})}$$

$$= \frac{\left(\prod_{i=1}^{N} \frac{1}{\sqrt{2\pi}\sigma_n}\right) \frac{1}{\sqrt{2\pi}\sigma_a}}{p(\mathbf{R})} \, exp\left\{ -\frac{1}{2}\left[\frac{\sum_{i=1}^{N}(R_i - A)^2}{\sigma_n^2} + \frac{A^2}{\sigma_a^2} \right] \right\}$$

<u>After rearranging</u>.........................

$$p(A|\mathbf{R}) = k(\mathbf{R})\, exp\left\{ -\frac{1}{2\sigma_p^2}\left[A - \frac{\sigma_a^2}{\sigma_a^2 + \sigma_n^2/N}\left(\sum_{i=1}^{N}\frac{R_i}{N}\right)\right]^2 \right\}$$

where $\sigma_p^2 = \left(\frac{1}{\sigma_a^2} + \frac{N}{\sigma_n^2}\right)^{-1} = \frac{\sigma_a^2 \sigma_n^2}{N\sigma_a^2 + \sigma_n^2}$ - a posteriori variance

We see that $p(A|\mathbf{R})$ is just a Gaussian density

$\Rightarrow$ $$\hat{a}_{ms}(\mathbf{R}) = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_n^2/N}\left(\frac{1}{N}\sum_{i=1}^{N} R_i\right)$$

Because the a posteriori variance is not a function of **R**, $\mathcal{R}_{ms} = \sigma_p^2$.

- $l(\mathbf{R}) = \sum_{i=1}^{N} R_i$ is a <u>sufficient statistic</u>.

  just like in the case of detection problem.

- Because the density is Gaussian the maximum value of the $p(A|\mathbf{R})$ occurs at the conditional mean.
  $\therefore\ \hat{a}_{map}(\mathbf{R}) = \hat{a}_{ms}(\mathbf{R})$

- Because the conditional median of a Gaussian density occurs at the conditional mean,
  $\hat{a}_{abs}(\mathbf{R}) = \hat{a}_{ms}(\mathbf{R})$

$\Rightarrow$ In this particular case all three cost function lead to the same estimate.

- This invariance to the choice of a cost function is a useful feature as the decision about $C(a_\in)$ is often subjective.

* **Real (Non Random) Parameter Estimation**

A – unknown but not random
$p(\mathbf{R}|A)$ given.

consider a Modified Bayes procedure – MSE criterion.

$$\mathcal{R}(A) \cong \int_{-\infty}^{\infty} [\hat{a}(\mathbf{R}) - A]^2 \, p(\mathbf{R}\,|\,A) d\mathbf{R}$$

$\Rightarrow$ $\hat{a}_{ms}(\mathbf{R})$ = A minimizing $\mathcal{R}$ (A)
Since A is unknown, this results is not useful.
Thus, consider

$$E[\hat{a}(\mathbf{R})] = \int_{-\infty}^{\infty} \hat{a}(\mathbf{R}) \, p(\mathbf{R}\,|\,A) d\mathbf{R}$$

the expected value of the estimate

If $E[\hat{a}(\mathbf{R})] = A$ estimate "unbiased"
$= A + B$ where B is not a function of A – known bias
$= A + B(A)$ - unknown bias

A second measure of quality of the estimate is

$$\text{var}\,[\hat{a}(\mathbf{R})-A] = E\{[\hat{a}(\mathbf{R})-A]^2\} - B^2(A)$$

Consider

$$\hat{a}(\mathbf{R}) = \arg\max_{A} p(\mathbf{R}|A)$$

$\longrightarrow$ likelihood function

$\Rightarrow$ Maximum likelihood estimate
$\hat{a}_{ml}(\mathbf{R})$.

MLE equation

$$\left.\frac{\partial \ln p(\mathbf{R}\,|\,A)}{\partial A}\right|_{A=\hat{a}_{ml}(\mathbf{R})} = 0$$

ML says nothing about a priori information; similar to map(MAP) if in MAP $p(A) = k$.

Example

$y = x + n \longrightarrow$ zero mean r.v.

$\Rightarrow \quad \hat{x}(y) = y$

- $r_i = A + n_i$, $i = 1, 2, ..., N$

$$p(\mathbf{R}|A) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi}\sigma_n} exp\left[-\frac{(R_i - A)^2}{2\sigma_n^2}\right]$$

$$\frac{\partial\, ln\ p(\mathbf{R}\,|\,A)}{\partial A} = \frac{N}{\sigma_n^2}\left(\frac{1}{N}\sum_{i=1}^{N} R_i - A\right)$$

$$\Rightarrow \quad \hat{a}_{ml}(\mathbf{R}) = \frac{1}{N}\sum_{i=1}^{N} R_i$$

<u>To find bias</u>

$$E[\hat{a}_{ml}(\mathbf{R})] = \frac{1}{N}\sum_{i=1}^{N} E(R_i) = \frac{1}{N}\sum_{i=1}^{N} A = A$$

so $\hat{a}_{ml}(\mathbf{R})$ is unbiased.

$$var[\hat{a}_{ml}(\mathbf{R}) - A] = \frac{\sigma_n^2}{N}$$

<u>CRAMER – RAO BOUND</u>

To consider the variance of any estimate $\hat{a}(\mathbf{R})$.

Thus,

a) If $\hat{a}(\mathbf{R})$ is any unbiased estimate of A. Then

$$var[\hat{a}(\mathbf{R}) - A] \geq \left(E\left\{\left[\frac{\partial\, ln\ p(\mathbf{R}\,|\,A)}{\partial A}\right]^2\right\}\right)^{-1}$$

or equivalenly

b) $$var[\hat{a}(\mathbf{R}) - A] \geq \left(-E\left[\frac{\partial^2\, ln\ p(\mathbf{R}\,|\,A)}{\partial A^2}\right]\right)^{-1}$$

where

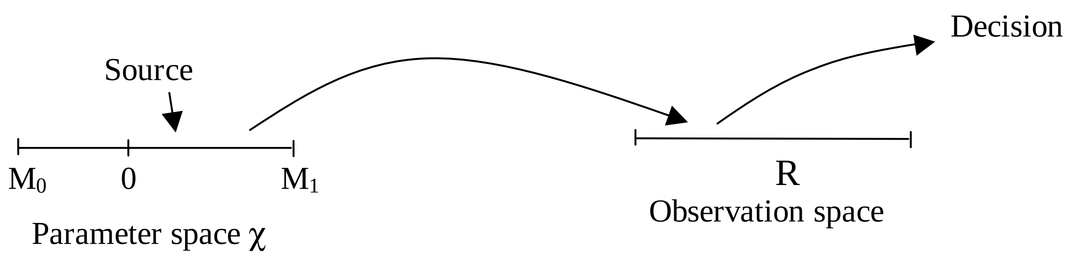
c) $\dfrac{\partial\, ln\ p(\mathbf{R}\,|\,A)}{\partial A}, \dfrac{\partial^2\, ln\ p(\mathbf{R}\,|\,A)}{\partial A^2}$ exist and absolutely integrable.

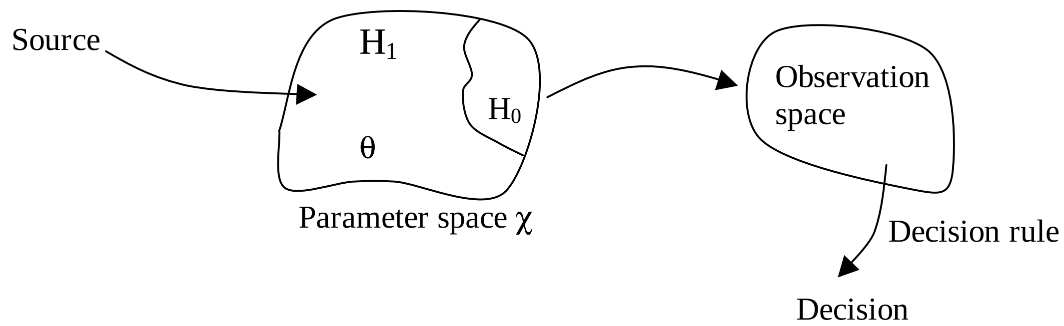
- Any estimate that satisfies the bound with an equality is called an "<u>efficient</u>" estimate.

⇒Note : If ML estimate is unbiased.

⇒ It is efficient.

Composite Hypothesis





Example

$$H_0 : p(R|H_0) = \frac{1}{\sqrt{2p}s} exp\left(-\frac{R^2}{2s^2}\right)$$

$$H_1 : p(R|H_1, M) = \frac{1}{\sqrt{2p}s} exp\left[-\frac{(R-M)^2}{2s^2}\right], \; M_0 \le M \le M_1$$

composite due to this

$M = 0 \Rightarrow H_0$

Assume $p(\mathbf{R}|M)$ is known for all values of M in χ.

<u>parameter</u> θ --- r.v. $\Rightarrow p(\theta|H_0), p(\theta|H_1)$

$$\Lambda(\mathbf{R}) \cong \frac{p(\mathbf{R}|H_1)}{p(\mathbf{R}|H_0)} = \frac{\int_c p(\mathbf{R}|q, H_1)\, p(q|H_1) dq}{\int_c p(\mathbf{R}|q, H_0)\, p(q|H_0) dq}$$

* known probability density on θ.

Assume $p(\mathrm{M}|\mathrm{H}_1) = \dfrac{1}{\sqrt{2\pi}\sigma_m} exp\left(-\dfrac{\mathrm{M}^2}{2\sigma_m^2}\right)$

Thus

$$\Lambda(\mathbf{R}) = \frac{\displaystyle\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} exp\left[-\frac{(\mathrm{R}-\mathrm{M})^2}{2\sigma^2}\right] . \frac{1}{\sqrt{2\pi}\sigma_m} exp\left[-\frac{\mathrm{M}^2}{2\sigma_m^2}\right] d\mathrm{M}}{\dfrac{1}{\sqrt{2\pi}\sigma} exp\left[-\dfrac{\mathrm{R}^2}{2\sigma^2}\right]}$$

Integrating and taking *ln*

$$\Rightarrow \quad \mathrm{R}^2 \underset{\mathrm{H}_0}{\overset{\mathrm{H}_1}{\gtrless}} \frac{2\sigma^2(\sigma^2+\sigma_m^2)}{\sigma_m^2}\left[ln\eta + \frac{1}{2} ln\left(1+\frac{\sigma_m^2}{\sigma^2}\right)\right]$$

- θ a r.v with unknown density ??
  minimax approach

or try several densities based on partial knowledge of that is available.

Hopefully test eventually will be insensitive to lack of exact knowledge in p.d.f.

- <u>θ - non random variable</u>

As θ has no probability density to average

⇒ Neyman – Pearson Tests

θ = M

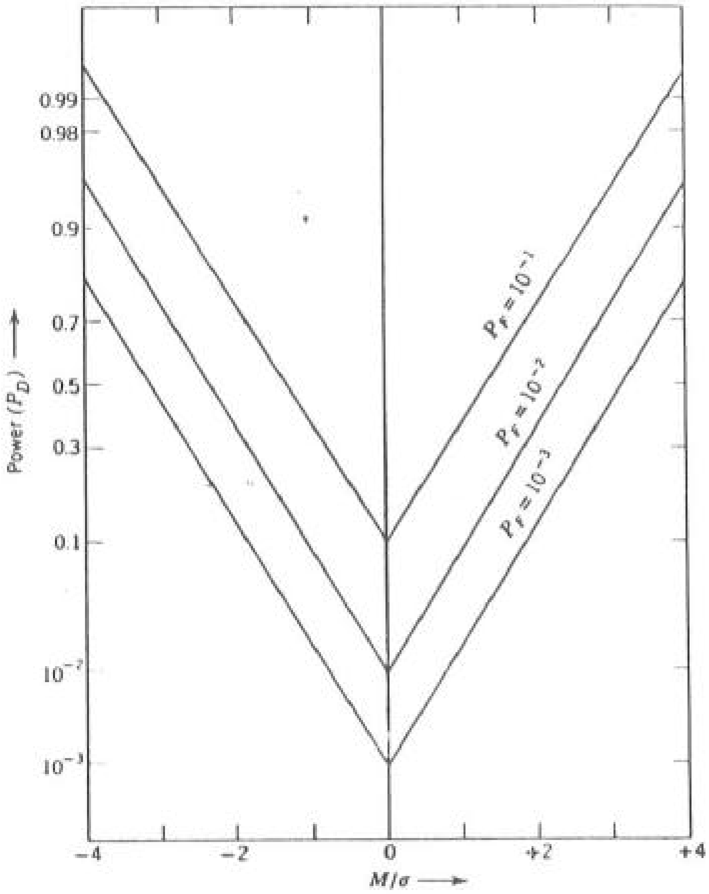
- perfect measurement bound.

Consider $\mathrm{H}_1 : p(\mathrm{R}|\mathrm{M}) = \dfrac{1}{\sqrt{2\pi}\sigma} exp\left[-\dfrac{(\mathrm{R}-\mathrm{M})^2}{2\sigma^2}\right]$, $\mathrm{M}_0 \le \mathrm{M} \le \mathrm{M}_1$

$\mathrm{H}_0 : p(\mathrm{R}|\mathrm{H}_0) = \dfrac{1}{\sqrt{2\pi}\sigma} exp\left(-\dfrac{\mathrm{R}^2}{2\sigma^2}\right)$

M – unknown non random parameter

- test can never be better than a hypothetical test where receiver knows M perfectly and then design the optimum LRT.
- ROC of any test ca be found by the ROC of this “perfect measurement test.”



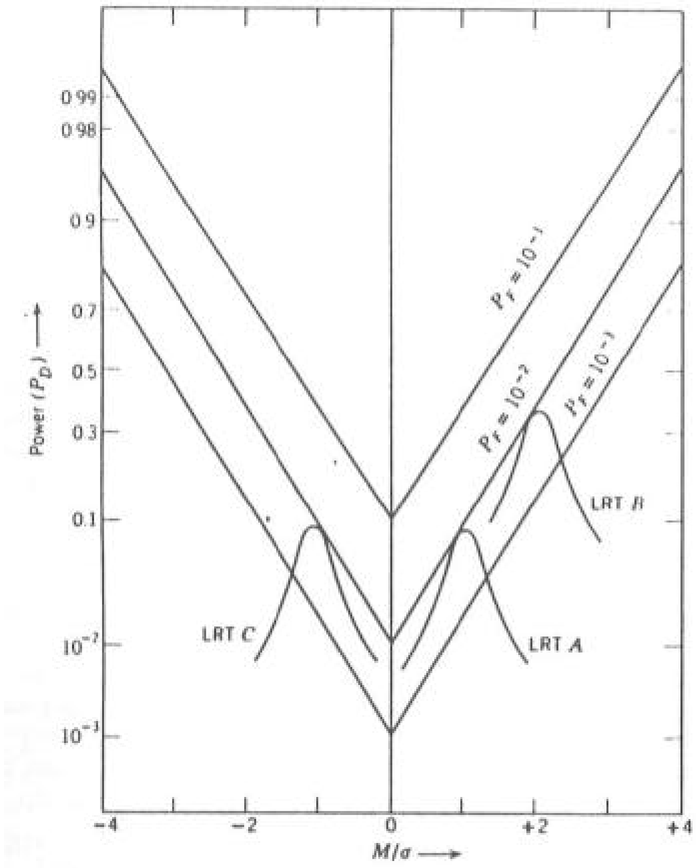
Power function for a perfect measurement test.

$H_0 = H_1$ for $M = 0$
$P_D = P_F$

- curves represent a bound on how well any test could do.
- The best performance if actual test curves equal the bound $\forall\ M \in \chi$. Such tests

$\Rightarrow$ "Uniformly most powerful" UMP.

- For a given P_F a UMP test has a $P_D \geq$ to any other test for $\forall\ M \in \chi$.
- First construct perfect measurement bound.

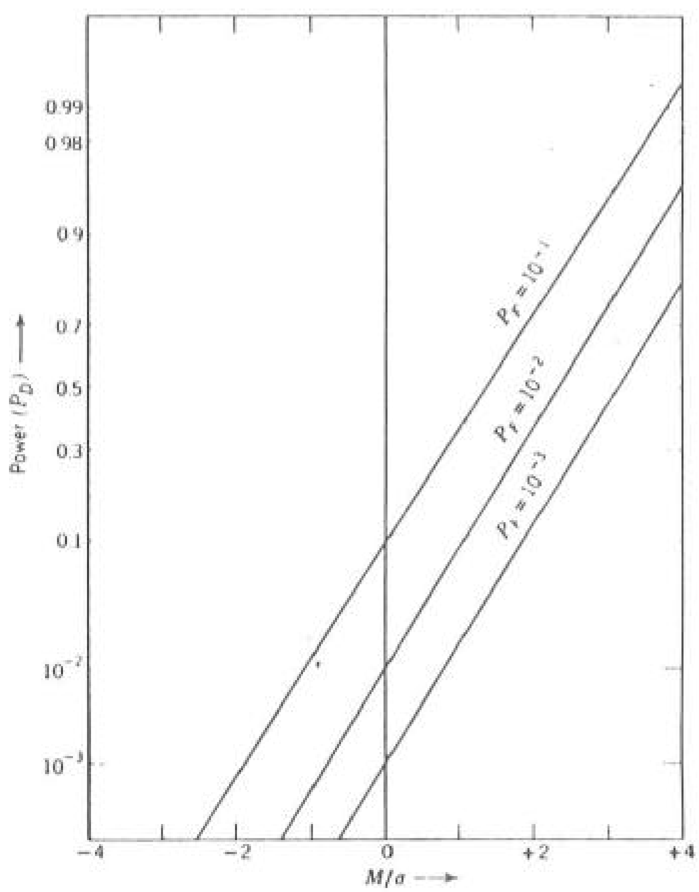


- LRT A
  B
  C

tests designed under the assumption M = 1, M = 2 M = -2

In each, P_D equals the bound at design points.

- For other values of M, the power of tests (P_D's) may or may not equal the bound.



Performance of LRT assuming (+)ve M. (correct curves)

- for (-)ve values of M; $P_D < P_F$

- It is clear that in general the bound can be reached for any particular θ ⇒ design LRT for that θ. A UMP test must be as good as any other test ∀θ.

<u>Property</u> : A UMP test exists if and only if the LRT for every $\theta \in \chi$ can be completely defined (including threshold) without knowledge of θ.

<u>Now in the example considered</u>

<u>test assume M>0</u>

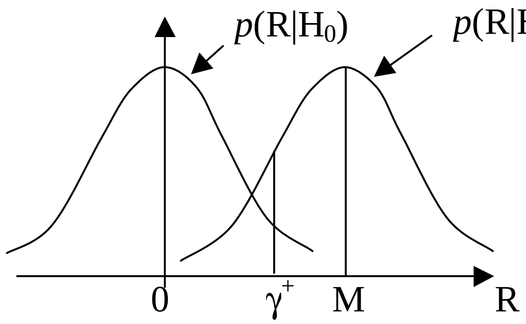
$$\text{LRT} \Rightarrow \qquad R \underset{H_0}{\overset{H_1}{\gtrless}} \gamma^+$$

$$P_F = \int_{\gamma^+}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} exp\left(-\frac{R^2}{2\sigma^2}\right) dR \text{ if } M>0$$

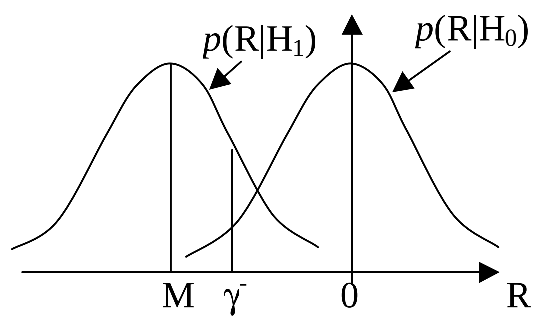
Similarly for M<0

$$\text{LRT} \Rightarrow \qquad R \underset{H_1}{\overset{H_0}{\gtrless}} \gamma^-$$

$$P_F = \int_{-\infty}^{\gamma^-} \frac{1}{\sqrt{2\pi}\sigma} exp\left(-\frac{R^2}{2\sigma^2}\right) dR \text{ if } M<0$$



M>0



M<0

- $M \geq 0 \Rightarrow$ UMP test exists. ($M_0 \geq 0$)
- $M \leq 0 \Rightarrow$ UMP test exists. ($M_1 \leq 0$)
- If $M_0 < 0$, $M_1 > 0$

  ⇒ UMP test does not exist.

Cases where UMP test does not exist

⇒ perfect measurement bound suggests

- * estimate θ assuming H_1 is true then
  estimate θ assuming H_0 is true

⇒ "Generalized likelihood ratio test"

$$\Lambda_g(\mathbf{R}) = \frac{\max_{\theta_1} p(\mathbf{R}|\theta_1)}{\max_{\theta_0} p(\mathbf{R}|\theta_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \gamma$$

where θ_1 ranges over all θ in H_1
θ_0 ranges over all θ in H_0

i.e. make a ML estimate of assuming H_1 true
evaluate $p(\mathbf{R}|\theta_1)$ for $\theta_1 = \hat{\theta}_1$
Similarly for $p(\mathbf{R}|\theta_0)$ for $\theta_0 = \hat{\theta}_0$

<u>Example</u>
Consider

$$p(\mathbf{R}|M, H_1) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi}\sigma} exp\left(-\frac{(R_i - M)^2}{2\sigma^2}\right)$$

- Composite Hypothesis θ = M

$$\hat{M}_1 = \frac{1}{N}\sum_{i=1}^{N} R_i \text{ - ML estimate}$$

$$p(\mathbf{R}|M, H_0) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi}\sigma} exp\left(-\frac{R_i^2}{2\sigma^2}\right)$$

- a simple hypothesis.

⇒ $\Lambda_g(\mathbf{R}) \Rightarrow P_F, P_D$ – Power function.

Usually <u>UMP or GLRT will give satisfactory results</u>.

** Comment on multiple parameter estimation

⇒ Most of the ideas can be extended from the single parameter case

⇒ Estimate more than one parameter

Read 2.4.3 Van Trees

The General Gaussian Problem

⇒ Cases where conditional density of R is Gaussian.

Definition: A set of random variables $r_1, r_2, ..., r_N$ are defined as jointly Gaussian if all their linear combinations are Gaussian random variables.

Definition: A vector **r** is a Gaussian random vector when its components $r_1, r_2, ..., r_N$ are jointly Gaussian variables.

⇒ If $z = \sum_{i=1}^{N} g_i r_i \cong \mathbf{G}^T \mathbf{r}$

is a Gaussian random variable for all finite G^T, then **r** is a Gaussian vector.

⇒ $E[\mathbf{r}] = \mathbf{m}$

$\mathrm{cov}[\mathbf{r}] = E[(\mathbf{r} - \mathbf{m})(\mathbf{r}^T - \mathbf{m}^T)] \cong \Lambda$

If Λ is nonsingular

$$p(\mathbf{R}) = \left[(2p)^{N/2} |\Lambda|^{1/2}\right]^{-1} \exp\left[-\frac{1}{2}\left(\mathbf{R}^T - \mathbf{m}^T\right)\Lambda^{-1}(\mathbf{R} - \mathbf{m})\right]$$

Definition: A hypothesis testing problem is called a General Gaussian if $p(\mathbf{R}|H_i)$ is a Gaussian density on all hypotheses.

Similarly, estimation problem ⇒ $p(\mathbf{R}|\mathbf{A})$ Gaussian density $\forall$ A

⇒ General Gaussian case.

$$\mathbf{r} = \begin{bmatrix} r_1 \\ r_2 \\ . \\ . \\ r_N \end{bmatrix}, \; E[\mathbf{r}|H_1] = \begin{bmatrix} E(r_1 \mid H_1) \\ E(r_2 \mid H_1) \\ . \\ . \\ E(r_N \mid H_1) \end{bmatrix} \cong \begin{bmatrix} m_{11} \\ m_{12} \\ . \\ . \\ m_{1N} \end{bmatrix} \cong \mathbf{m}_1$$

The covariance matrix is

$$\mathbf{K}_1 \cong E[(\mathbf{r} - \mathbf{m}_1)\,(\mathbf{r}^T - \mathbf{m}_1^T)|H_1] = \begin{bmatrix} {}_1K_{11} & {}_1K_{12} & {}_1K_{13}\ldots & {}_1K_{1N} \\ {}_1K_{21} & {}_1K_{22} & {}_1K_{23}\ldots & {}_1K_{2N} \\ . & . & & . \\ . & . & & . \\ {}_1K_{N1} & & & {}_1K_{NN} \end{bmatrix}$$

under H_1

If $\mathbf{Q}_1 \cong \mathbf{K}_1^{-1} \Rightarrow \mathbf{Q}_1\mathbf{K}_1 = \mathbf{K}_1\mathbf{Q}_1 = \mathbf{I}$

Thus, $p(\mathbf{R}|H_1) = \left[(2p)^{N/2}|\mathbf{K}_1|^{1/2}\right]^{-1} \exp\left[-\frac{1}{2}(\mathbf{R}^T - \mathbf{m}_1^T)\mathbf{Q}_1(\mathbf{R} - \mathbf{m}_1)\right]$

Similarly,

$$p(\mathbf{R}|H_0) = \left[(2p)^{N/2}|\mathbf{K}_0|^{1/2}\right]^{-1} \exp\left[-\frac{1}{2}(\mathbf{R}^T - \mathbf{m}_0^T)\mathbf{Q}_0(\mathbf{R} - \mathbf{m}_0)\right]$$

The LRT

$$\Rightarrow \quad \Lambda(\mathbf{R}) = \frac{|\mathbf{K}_0|^{1/2}\, exp\left[-\frac{1}{2}(\mathbf{R}^T - \mathbf{m}_1^T)\mathbf{Q}_1(\mathbf{R} - \mathbf{m}_1)\right]}{|\mathbf{K}_1|^{1/2}\, exp\left[-\frac{1}{2}(\mathbf{R}^T - \mathbf{m}_0^T)\mathbf{Q}_0(\mathbf{R} - \mathbf{m}_0)\right]} \underset{H_0}{\overset{H_1}{\gtrless}} \eta$$

Taking logarithm$\Rightarrow$

$$\frac{1}{2}(\mathbf{R}^T - \mathbf{m}_0^T)\mathbf{Q}_0(\mathbf{R} - \mathbf{m}_0) - \frac{1}{2}(\mathbf{R}^T - \mathbf{m}_1^T)\mathbf{Q}_1(\mathbf{R} - \mathbf{m}_1) \underset{H_0}{\overset{H_1}{\gtrless}} \ln \eta + \tfrac{1}{2}\ln|\mathbf{K}_1| - \tfrac{1}{2}\ln|\mathbf{K}_0| \cong \gamma^*$$

L.H.S difference of quadratic forms.

Special Cases

1. <u>Equal Covariance Matrices</u>

$\mathbf{K}_1 = \mathbf{K}_0 \cong \mathbf{K}$ $\qquad$ $\mathbf{m}_1 \neq \mathbf{m}_0$

$\mathbf{Q} = \mathbf{K}^{-1} \rightarrow$ symmetric

The above simplifies to

$$\left(\mathbf{m}_1^T - \mathbf{m}_0^T\right)\mathbf{Q}\,\mathbf{R} \underset{H_0}{\overset{H_1}{\gtrless}} \ln\eta + \frac{1}{2}\left(\mathbf{m}_1^T\mathbf{Q}\mathbf{m}_1 - \mathbf{m}_0^T\mathbf{Q}\mathbf{m}_0\right) \cong \gamma_*$$

Let $\Delta\mathbf{m} = \mathbf{m}_1 - \mathbf{m}_0$

$$\Rightarrow l(\mathbf{R}) \cong \Delta\mathbf{m}^T\mathbf{Q}\mathbf{R} \underset{H_0}{\overset{H_1}{\gtrless}} \gamma_*$$

or

$$l(\mathbf{R}) \cong \mathbf{R}^T\mathbf{Q}\Delta\mathbf{m} \underset{H_0}{\overset{H_1}{\gtrless}} \gamma_*$$

<u>Scalar</u> Gaussian r.v $\Leftarrow$ a linear transformation byGaussian r.v.s.

Thus, performance can be completely characterized by

$$d^2 \cong \frac{\left[E(l \mid H_1) - E(l \mid H_0)\right]^2}{var(l \mid H_0)} \longleftarrow \text{Normalizing}$$

Thus

$E(l|H_1) = \Delta\mathbf{m}^T\mathbf{Q}\mathbf{m}_1$

$E(l|H_0) = \Delta\mathbf{m}^T\mathbf{Q}\mathbf{m}_0$

$var(l|H_0) = E\{[\Delta\mathbf{m}^T\mathbf{Q}(\mathbf{R}-\mathbf{m}_0)]\,[(\mathbf{R}-\mathbf{m}_0^T)\,\mathbf{Q}\Delta\mathbf{m}]\}$

$= \Delta\mathbf{m}^T\mathbf{Q}\Delta\mathbf{m}$ $\qquad$ (because $E\{(\mathbf{R}-\mathbf{m}_0)(\mathbf{R}-\mathbf{m}_0^{\;T})\} = \mathbf{K} = \mathbf{Q}^{-1}$)

Therefore

$$d^2 = \Delta\mathbf{m}^T\mathbf{Q}\Delta\mathbf{m}$$

If $\mathbf{K} = \sigma^2 \mathbf{I} \Rightarrow \mathbf{Q} = \frac{1}{s^2}\mathbf{I}$

$$d^2 = \Delta\mathbf{m}^T \frac{1}{s^2}\mathbf{I}\,\Delta\mathbf{m} = \frac{1}{s^2}\Delta\mathbf{m}^T\mathbf{I}\,\Delta\mathbf{m}$$

$$= \frac{1}{s^2}|\Delta\mathbf{m}|^T = \frac{1}{s^2}[(m_{11} - m_{01})^2 + (m_{12} - m_{02})^2 + \ldots]$$

$$l(\mathbf{R}) = \frac{1}{s^2}\Delta\mathbf{m}^T\mathbf{R}$$

The sufficient statistic is just the dot (scalar) product between **R** and the mean difference vector **Δm**.

- <u>Equal mean vectors</u>

$m_1 = m_0 \cong \mathbf{m} \equiv 0$ without loss of generality

$$\Delta\mathbf{Q} = \mathbf{Q}_0 - \mathbf{Q}_1$$

$$l(\mathbf{R}) \cong \mathbf{R}\Delta\mathbf{Q}\mathbf{R} \quad \underset{H_0}{\overset{H_1}{\gtrless}} \quad \gamma$$

$\mathbf{R}^{\perp}$, $\Delta\mathbf{Q}\mathbf{R} \Rightarrow l(\mathbf{R})$ is not a Gaussian r.v.

Orthogonal Representation of Random Processes

First, consider deterministic problem.

- Review : Orthogonal representation of signals and white noise.

  ↓

  $x(t)$ – [0, T] $\qquad E_x = \int_0^T x^2(t)dt < \infty$

To specify x(t) by a countable set of numbers

$\Rightarrow \qquad x(t) = \sum_{i=1}^{\infty} x_i f_i(t)$

↘ some set of orthogonal functions

Questions :

1) Only finite number N practical, how should we choose the coefficients to minimize the mean-square approximation error?

2) N ↑ error → 0; when ?

*) $\qquad C_N(t) = x(t) - \sum_{i=1}^{N} x_i f_i(t)$

$$E_e(N) \cong \int_0^T e_N^2(t)dt = \int_0^T \left[x(t) - \sum_{i=1}^{N} x_i f_i(t) \right]^2 dt$$

↓

minimize ∀ N

Differentiate w.r.t particular $x_j \Rightarrow 2\int_0^T \left[x(t) - \sum_{i=1}^{N} x_i f_i(t) \right](-f_j(t))dt$

$\Rightarrow x_j = \int_0^T x(t) f_j(t)dt$

2^{nd} derivative (+) ve ⇒ minimum

The choice of coefficient does not change as N is increased because of orthogonality of functions.

$$E_e(N) = E_x - 2\sum_{i=1}^{N} \int_0^T x(t)x_i f_i(t)dt + \int_0^T \sum_{i=1}^{N} \sum_{i=1}^{N} x_i x_j f_i(t) f_j(t)dt$$

$$= E_x - \sum_{i=1}^{N} x_i^2$$

$x_i^2 \geq 0$, the error is a monotonic – decreasing function of N.

If $\lim_{N\to\infty} E_e(N) = 0$

$\forall$ x(t) with finite energy $\Rightarrow$ $f_i(t)$ = 1, 2, ... are a "Complete orthonormal – (CON) set" over the interval [0, T], for the class of functions with finite energy.

- For CON sets $E_x = \sum_{i=1}^{\infty} x_i^2$
  - "Parseval's theorem" – x_i^2 represents the energy in a particular componetn of the signal.

Generation of x_i's $\rightarrow$ 1) integrate and dumps – correlation
2) Matched Filtered
$h_i(\tau) = \phi_i(T-\tau)$

- Random Process Characterization

$m_x(t) = E(x_t) = \int_{-\infty}^{\infty} x_t\, p(x_t) dx_t$ - Mean Value

Correlation

$$R_x(t, u) = \iint_{-\infty}^{+\infty} x_t x_u\, p(x_t, x_u) dx_t dx_u$$

Co-variance

$$K_x(t, u) = E\{[x_t - m_x(t)]\,[x_u - m_x(u)]\}$$
$$= R_x(t, u) - m_x(t)\, m_x(u)$$

$$K_x(t, u) = K_x(u, t) \text{ – symmetry.}$$

$f(t)$ – deterministic function with finite energy;

$$x_f = \int_0^T x(t)\, f(t) dt \Rightarrow \text{r.v.}$$

$$E[x_f] = \int_0^T m_x(t)\, f(t) dt$$

$$\text{var}(x_f) = E[(x_f - \bar{x}_f)^2]$$
$$= E\{\int_0^T [x(t) - m_x(t)]\, f(t) dt \int_0^T [x(u) - m_x(u)]\, f(u) du$$
$$= \iint_0^T f(t) K_x(t, u)\, f(u) dt du$$

$$\mathrm{var}(x_f) \ge 0$$

$$\Rightarrow \quad \int\!\!\int_0^T f(t)K_x(t,u)\, f(u)\,dt\,du \ \ge 0$$

If the inequality strict $\Rightarrow$ $K_x(t, u)$ (+) ve definite.

Now, for a scalar random process x(t).
Choose a CON set of deterministic functions

$$x(t) = \sum_{i=1}^{N} x_i\phi_i(t), \quad 0\le t\le T$$
$$\underset{lim}{N\to\infty}$$

$$x_i \cong \int_0^T x(t)\phi_i(t)dt$$

Here consider mean-square convergence;

$$x(t) = \lim_{N\to\infty} \sum_{i=1}^{N} x_i\, f_i(t), \quad 0\le t\le T$$

limit in the mean $\Rightarrow$ $\lim_{N\to\infty} E\left[\left(x_t - \sum_{i=1}^{N} x_i\, f_i(t)\right)^2\right] = 0, \ 0\le t\le T$

Find $\phi_i(t)$ that leads to "uncorrelated" coefficients

$\Rightarrow$ If $E(x_i) = m_i$ want
$E[(x_i - m_i)(x_j - m_j)] = \lambda_i\delta_{ij}$

Assume $m_i = 0$

$E(x_i) = 0$, $E(x_i\, x_j) = 0$, $i\neq j$
$E(x_i^2) = \lambda_i$

* Expected value of energy x_i^2 along $\phi_i(t) \Rightarrow \lambda_i \ge 0 \ \ \forall\, i$
If $K_x(t, u)$ (+) ve definite $\lambda_i>0$, $\forall$ i, otherwise <u>at least one</u> $\lambda_i = 0$.

$$E(x_ix_j) = E\left[\int_0^T x(t)\phi_i(t)dt\int_0^T x(u)\phi_j(u)du\right]$$

$$\lambda_i\delta_{ij} = \int_0^T \phi_i(t)dt\int_0^T K_x(t,u)\phi_j(u)du \qquad \forall\, i,j$$

$$\Rightarrow \quad \int_0^T K_x(t,u)\phi_j(u)du = \lambda_i\phi_j(t) \qquad , 0\le t\le T$$

$\phi_i(t)$ – eigen functions λ_i – eigen values.

- <u>Homogeneous Integral Equation</u> (HIE)

The series expansion developed

⇒ "Karhunen – Loeve" expansion

provides 2nd moment characterization of uncorrelated random variables.

Properties :

1) ? at least one $\phi(t)$ and real number $\lambda \neq 0$ that satisfy HIE.
2) Can always normalize $\phi_i(t)$.
3) $\phi_1(t), \phi_2(t) \rightarrow \lambda \Rightarrow C_1\phi_1(t) + C_2\phi_2(t) \Rightarrow \lambda$
4) $\lambda_n \neq \lambda_m \quad < \phi_n, \phi_m > = 0$
5) For any λ, there is at most a finite number of l.i. eigen functions. l.i. in algebraic sense.
   These can always be orthonormalized.
   ⇒ Gram – Schmidt Procedure
6) $K_x(t, u) = \sum_{i=1}^{\infty} \lambda_i \phi_i(t)\phi_i(u), \quad 0 \le t,u \le T$
   - Mercers Theorem
7) If $K_x(t, u)$ (+) ve definite
   the eigen functions form a CON.
8) If not (+) ve definite i.e. (+) ve semi-definite
   ⇒ along some directions energy = 0

   To get a CON set need to augment with $\phi_i(t)$ corresponding to zero eigen values.

9) $E\left\{\int_0^T x^2(t)dt\right\} = \int_0^T K_x(t,t)dt = \sum_{i=1}^{\infty} \lambda_i$

Example

Let the random process be white noise. Then $K_x(t, u) = \sigma^2\delta(t-u)$

Integral equation

$$\sigma^2 \int_0^T \delta(t-u)\phi(u)du = \sigma^2\phi(t)$$

(with σ^2 on the right-hand side marked by an arrow as λ)

⇒ Any CON set satisfies the integral equation and has the property that

$$\sum_{i=1}^{\infty} \phi_i(t)\phi_j(u) = \delta(t-u)$$

Example Winer Process – (To model Brownian Motion)

$$K_x(t, u) = \sigma^2 \min(u,t) = \begin{cases} \sigma^2 u, & u \le t \\ \sigma^2 t, & t \le u \end{cases}$$

$x(0) = 0 \qquad E[x(t)] = 0$

$E[x^2(t)] = \sigma^2 t \qquad p(x_t) = \dfrac{1}{\sqrt{2\pi\sigma^2 t}} \, exp\left(-\dfrac{x_t^2}{2\sigma^2 t}\right)$

If $t_3 > t_2 > t_1$

$\Rightarrow \quad \left(x_{t_3} - x_{t_2}\right), \left(x_{t_2} - x_{t_1}\right)$ are statistically independent.

?? * process defined for $t \ge 0$.

The integral equation

$$\lambda\phi(t) = \int_0^T K_x(t,u)\phi(u)du \qquad , 0 \le t \le T$$

$$= \sigma^2 \int_0^t u\phi(u)du + \sigma^2 t \int_t^T \phi(u)du$$

Differentiate w.r.t time t

$$\lambda\dot{\phi}(t) = \sigma^2 t\,\phi(t) + \sigma^2 \int_t^T \phi(u)du - \sigma^2 t\,\phi(t)$$

$$= \sigma^2 \int_t^T \phi(u)du$$

$$\lambda\ddot{\phi}(t) = -\sigma^2 \phi(t) \Rightarrow \ddot{\phi}(t) + \frac{\sigma^2}{\lambda}\phi(t) = 0$$

$\lambda > 0$ Test solution $\rightarrow \phi_\lambda(t) = A_n \cos\left(\dfrac{\sigma}{\sqrt{\lambda}} t\right) + B_n \sin\left(\dfrac{\sigma}{\sqrt{\lambda}} t\right)$

Substitute back into the integral equation

Choose B_n to normalize the solution $A_n = 0$, $\lambda_n = \dfrac{\sigma^2 T^2}{\left(n - 1/2\right)^2 \pi^2}$, n = 1, 2,

$$\therefore \phi_n(t) = \sqrt{\frac{2}{T}}\, sin\left[\left(n - 1/2\right)\frac{\pi t}{T}\right]$$

Detection in White Gaussian Noise

$H_i : R(t) = s_i(t) + W(t)\ ;\ 0 \le t \le T$

- Review *(a) Simple Binary Problem

$H_1 : R(t) = \sqrt{E}\, s_0(t) + w(t)$
$H_0 : R(t) = w(t)$

$$\begin{cases} 0 \le t \le T \\ \int_0^T s_0^2(t)dt = 1 \end{cases}$$

$w(t) \Rightarrow N_0/2$ watts/Hz psdf

K-L expansion

$\phi_1(t) = s_0(t)$

.
.
.

$$\begin{cases} \text{orthonormal} \\ \text{to} \quad s_0(t) \end{cases}$$

$R_i = \int_0^T R(t)\phi_i(t)dt = \sqrt{E}\,\delta_{ij} + w_i$ under H_1

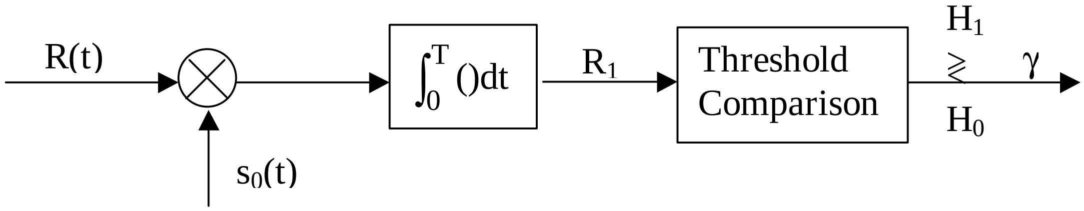
$= w_i$ under H_0

- R_i's are statistical independent.
  R_i for $i > 1$ are the same for both hypothesis.

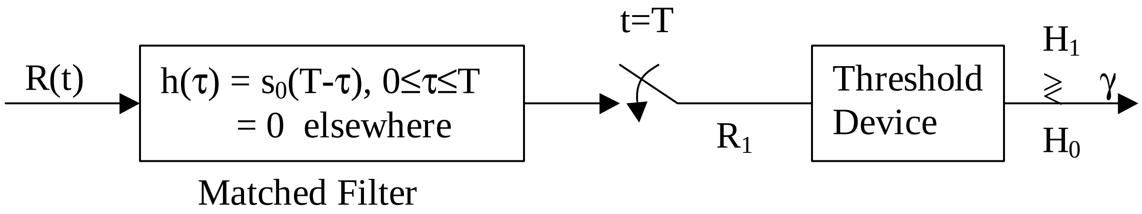
$\Rightarrow$ R_1 is a sufficient statistic
$\Rightarrow$ The optimum test is

$$R_1 \underset{H_0}{\overset{H_1}{\gtrless}} \text{threshold} = \gamma$$

Receiver (Demodulator)



Correlation Receiver



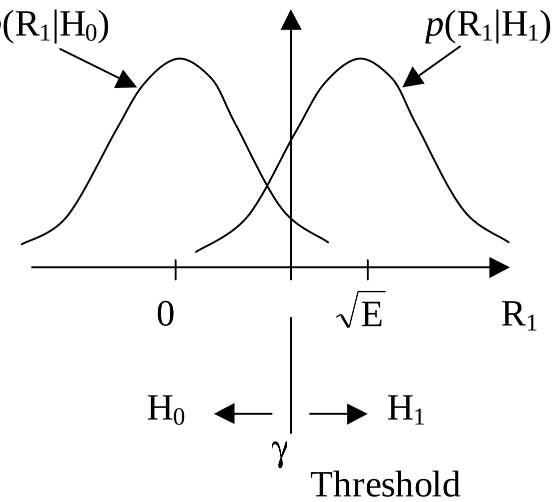
<u>Performance</u>

$$P_F = Q\left(\frac{\gamma}{\sqrt{N_0/2}}\right)$$

$$P_D = 1 - Q\left(\frac{\sqrt{E}-\gamma}{\sqrt{N_0/2}}\right)$$

$$d^2 = \frac{\left[E(R_1 \mid H_1) - E(R_1 \mid H_0)\right]^2}{\sigma_{R_1|H_1} . \sigma_{R_1|H_0}}$$

$$= \frac{2E}{N_0} = \frac{E}{N_0/2}$$



- Performance only depends on $\sqrt{E}$, independent of signal shape - reason is because the noise is white.

(b) <u>General Binary Detection Problem</u>

$H_1 : R(t) = \sqrt{E_1}\, s_1(t) + w(t),$ $\quad 0 \le t \le T \quad w(t) \to N_0/2$

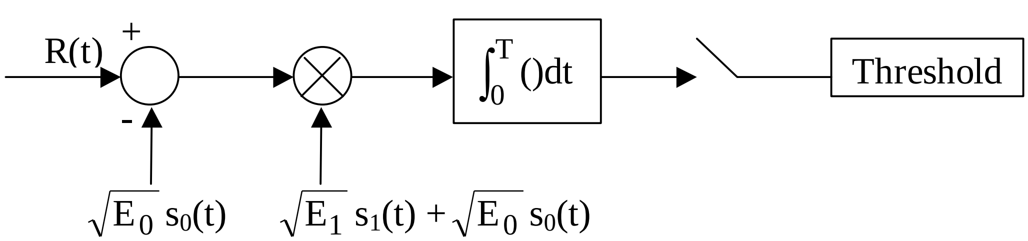
$H_0 : R(t) = \sqrt{E_0}\, s_0(t) + w(t)$ $\quad \int_0^T s_1^2(t)dt = 1$

<u>One solution</u> – consider $R'(t) = R(t) - \sqrt{E_0}\, s_0(t)$

$\Rightarrow$ $H_1 : R'(t) = (\sqrt{E_1}\, s_1(t) - \sqrt{E_0}\, s_0(t)) + w(t)$

$H_0 : R'(t) = w(t)$

<u>Thus optimum solution</u>



2nd Solution

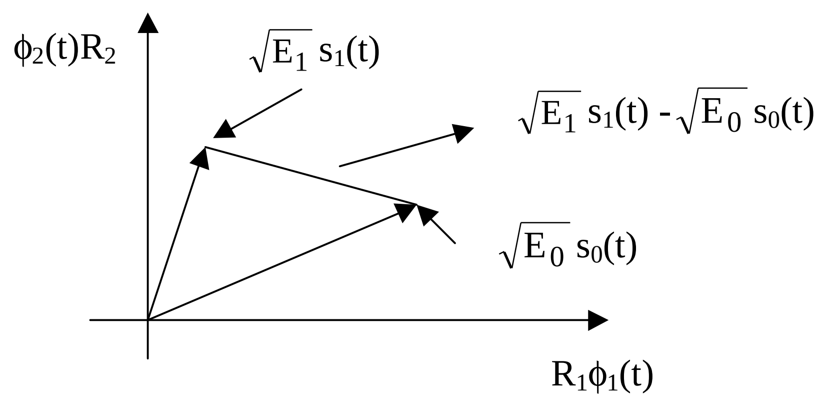
Let $p = \int_0^T s_0(t)s_1(t)dt$

CON set – so that $s_0(t)$, $s_1(t)$ are orthogonal to $\phi_3(t)$, $\phi_4(t)$, ...

Let $\phi_1(t) = s_0(t)$, $\phi_2(t) = \dfrac{s_1(t) - p s_0(t)}{\sqrt{1-p^2}}$

Gram – Schmidt Procedure

Project R(t) onto $\phi_i(t)$ – R_1, R_2, R_3, ...
R_i for $i \geq 3$ are indifferent.
Have a 2-D problem – sufficient statistics



After simplification test:

$$\int_0^T R(t)\left(\sqrt{E_1}s_1(t) - \sqrt{E_0}s_0(t)\right)dt \underset{H_0}{\overset{H_1}{\gtrless}} \frac{N_0}{2} \ln \eta + \frac{E_1 - E_0}{2}$$

Special case
$C_{ij} = 1 - \delta_{ij} \Rightarrow$ Minimum Error Probability
$\Pr[H_0] = \Pr[H_1]$

$\Rightarrow$ decision boundary is line bisector
$E_0 = E_1 \Rightarrow$ bisector goes through origin

$$\boxed{\int_0^T R(t)s_1(t)dt \underset{H_0}{\overset{H_1}{\gtrless}} \int_0^T R(t)s_0(t)dt}$$

⇒ "Largest of" Receiver

- Can change coordinates – orthonormal functions; so that only one $\phi(t)$ is needed. ⇒ check early notes.

⇒ One dimensional case, *l* Gaussian.

$$d^2 = \frac{2}{N_0}\left(E_1 + E_0 - 2\,p\sqrt{E_0 E_1}\right)$$

- <u>M – ary detection in WGN</u>

$H_1 : R(t) = \sqrt{E_1}\, s_1(t) + w(t)$, $\quad 0 \le t \le T$

$H_2 : R(t) = \sqrt{E_2}\, s_2(t) + w(t)$ $\quad \int_0^T s_i^2(t)dt = 1$

. $\quad \int_0^T s_i(t) s_j(t) dt = r_{ij}$

.

$H_M : R(t) = \sqrt{E_M}\, s_M(t) + w(t)$

$w(t) \to N_0/2$

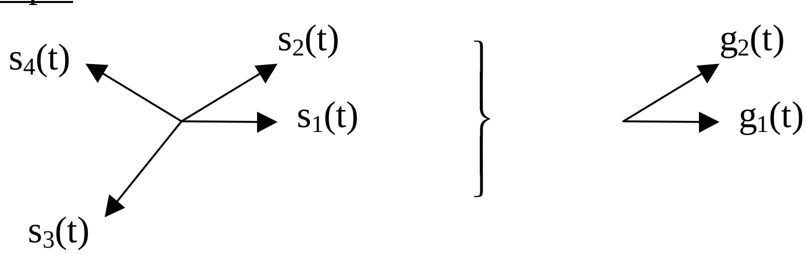
Let $g_1(t), g_2(t), \ldots, g_L(t)$ be such that

a) linearly independent

b) span $\{g_1(t), g_2(t), \ldots, g_L(t)\}$ = span$\{s_1(t), s_2(t), \ldots, s_M(t)\}$

<u>Note</u> : $L \le M$

<u>Example</u>



$g_i(t)$ are not necessarily orthogonal.

Let $R_i = \int_0^T R(t) g_i(t) dt$; i = 1, 2, ..., L

$\mathbf{R} = (R_1, R_2, \ldots, R_L)^T$ ....(*)

Claim **R** is a sufficient statistic.

Proof: Let $\phi_1(t), \phi_2(t), \ldots$ be a CON set. So that

$\text{span}\{\phi_1(t), \phi_2(t), \ldots, \phi_L(t)\} = \text{span}\{g_1(t), g_2(t), \ldots, g_L(t)\}$

(Gram- Schmidt procedure)

$$R(t) = \sum_{i=1}^{\infty}\left[\int_0^T R(t)\phi_i(t)dt\right]\phi_i(t) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(**)$$

Sufficient statistic is the first L of the coefficients.

For i = 1, 2, ..., L

$$\phi_i(t) = \sum_{i=1}^{L}\beta_{ij}g_j(t) \Rightarrow \int_0^T R(t)\phi_i(t)dt = \sum_{i=1}^{L}\beta_{ij}R_j$$

i.e, the first L coefficients of the expansion of R(t) in (**) are linear combination of the L coefficients R_i in (*).

$\Rightarrow$ $R_1, R_2, \ldots, R_L$ of (*) are sufficient statistics. QED

Let $\int_0^T g_i(t)g_j(t)dt = n_{ij}$

$H_i : R = \sqrt{E_i}\,\mathbf{s}_i + \mathbf{w}$

$$\mathbf{s}_i = \begin{bmatrix} s_{i1} \\ s_{i2} \\ . \\ . \\ s_{iL} \end{bmatrix}, \qquad s_{ij} = \int_0^T s_i(t)g_j(t)dt$$

$$\mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ . \\ . \\ w_L \end{bmatrix} = \begin{bmatrix} \int_0^T w(t)g_1(t)dt \\ \int_0^T w(t)g_2(t)dt \\ . \\ . \\ \int_0^T w(t)g_L(t)dt \end{bmatrix}, \quad \mathbf{R} = \begin{bmatrix} R_1 \\ R_2 \\ . \\ . \\ R_L \end{bmatrix}$$

$$E[w_i w_j] = E\left\{\int_0^T w(t)g_i(t)dt \int_0^T w(\tau)g_j(\tau)d\tau\right\}$$

$$= \int_0^T\int g_i(t)g_j(\tau)\frac{N_0}{2}\delta(t-\tau)dtd\tau$$

$$= \frac{N_0}{2}\int_0^T g_i(t)g_j(t)dt$$

$$= \frac{N_0}{2}n_{ij}$$

$\Rightarrow$ $R_1, R_2, \ldots$ are correlated in general.

Now consider orthonormal expression.

$\phi_1(t), \phi_2(t), \ldots, \phi_L(t) \Rightarrow \text{span}\{s_1(t), s_2(t), \ldots, s_M(t)\}$

<u>Gram – Schmidt Procedure</u>

$\phi_1(t) = s_1(t)$

$\phi_2(t) = (1- p_{12}^2)^2\, [s_2(t) - p_{12}s_1(t)]$

$\phi_3(t) = c_3\, [s_3(t) - c_1\phi_1(t) - c_2\phi_2(t)]$

normalize

c_1, c_2 – orthogonally

<u>Now</u>, $R_i = \int_0^T R(t)\phi_i(t)dt$

$$H_i : \mathbf{R} = \sqrt{E_i}\,\mathbf{s}_i + \mathbf{w}, \qquad \mathbf{R} = \begin{bmatrix} R_1 \\ R_2 \\ . \\ . \\ R_L \end{bmatrix}$$

$$\mathbf{s}_i = \begin{bmatrix} s_{i1} \\ s_{i2} \\ . \\ . \\ s_{iL} \end{bmatrix}, \qquad s_{ij} = \int_0^T s_i(t)\phi_j(t)dt$$

$$E[w_i w_j] = E\left\{\int_0^T w(t)\phi_i(t)dt \int_0^T w(\tau)\phi_j(\tau)d\tau\right\} = \frac{N_0}{2}\delta_{ij}$$

$E[\mathbf{w}_i\mathbf{w}_j^T] = \frac{N_0}{2} I$

<u>Likelihood Ratio Test</u> i, j

$$\Lambda_{ij}(\mathbf{R}) = \frac{p(\mathbf{R} \mid H_i)}{p(\mathbf{R} \mid H_j)}$$

$$= \frac{\left[(2\pi)^{L/2}\left|\frac{N_0}{2} I\right|^{\frac{1}{2}}\right]^{-1} exp\left\{-\frac{1}{2}(\mathbf{R}-\mathbf{s}_i)^T\left(\frac{N_0}{2} I\right)^{-1}(\mathbf{R}-\mathbf{s}_i)\right\}}{(i \to j)}$$

$$= \frac{exp\left\{-\frac{1}{N_0}\sum_{k=1}^{L}\left(R_k - \sqrt{E_i}s_{ik}\right)^2\right\}}{exp\left\{-\frac{1}{N_0}\sum_{k=1}^{L}\left(R_k - \sqrt{E_j}s_{jk}\right)^2\right\}}$$

$$\ln \Lambda_{ij}(\mathbf{R}) = \frac{1}{N_0}\left[\sum_{k=1}^{L}\left(R_k - \sqrt{E_j}s_{jk}\right)^2 - \sum_{k=1}^{L}\left(R_k - \sqrt{E_i}s_{ik}\right)^2\right]$$

$$= \frac{1}{N_0}\left[\left\|\mathbf{R}_k - \sqrt{E_j}\mathbf{s}_{jk}\right\|^2 - \left\|\mathbf{R}_k - \sqrt{E_i}\mathbf{s}_{ik}\right\|^2\right]$$

<u>After simplification</u>

$$= \frac{2}{N_0}\left[\sum_{k=1}^{L}\sqrt{E_j}\left(R_k s_{jk}\right) - \sum_{k=1}^{L}\sqrt{E_i}\left(R_k s_{ik}\right)\right]$$

$$+\frac{1}{N_0}\left[E_j\underbrace{\sum_{k=1}^{L}s_{jk}^2}_{=1} - E_i\underbrace{\sum_{k=1}^{L}s_{ik}^2}_{=1}\right]$$

$$= \int_0^T s_i^2(t)dt$$

$$\sum_{k=1}^{L} R_k s_{ik} = \int_0^T R(t)s_i(t)dt$$

Since $s_i(t) = \sum_{k=1}^{L} s_{ik}\phi_k(t)$ and $\int_0^T R(t)\phi_k(t)dt = R_k$

Therefore

$$\ln \Lambda_{ij}(\mathbf{R}) = \frac{2}{N_0}\left[\sqrt{E_i}\int_0^T R(t)s_i(t)dt - \sqrt{E_j}\int_0^T R(t)s_j(t)dt\right] + \frac{(E_j - E_i)}{N_0}$$

Consider MAP rule (min Pr[error]) → recall

Let $P_i = \Pr[H_i]$

$\therefore$ Compute $\Pr[H_i|\mathbf{R}] = \dfrac{p(\mathbf{R}\,|\,H_i).\Pr[H_i]}{p(\mathbf{R})}$, $\quad i = 1, 2, \ldots, M$

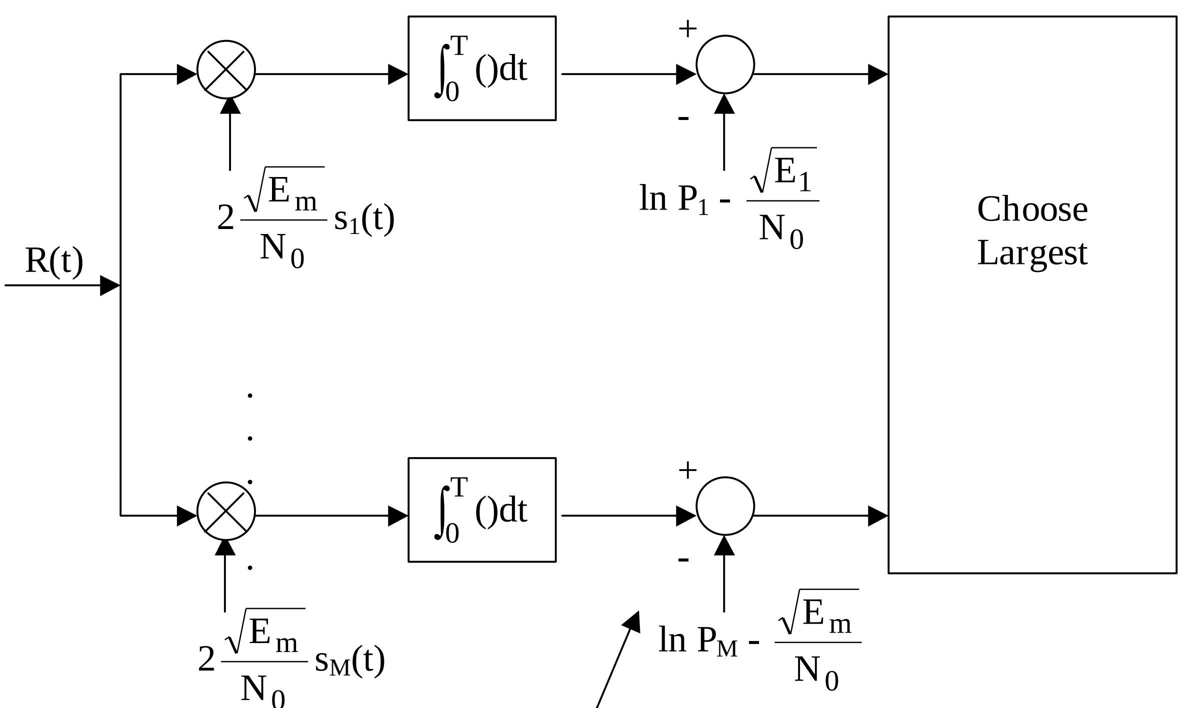
Choose largest

or choose largest of $\ln p(\mathbf{R}|H_i) + \ln P_i$

$\Rightarrow$ Choose largest of $\ln P_i - \dfrac{1}{N_0}\left[\sum_{k=1}^{L}\left(R_k - \sqrt{E_i}s_{ik}\right)^2\right]$

Throw away $\sum R_k^2$

$\Rightarrow$ Choose largest of $\ln P_i + \dfrac{2\sqrt{E_i}}{N_0}\int_0^T R(t)s_i(t)dt - \dfrac{E_i}{N_0}$



- Do not need to find an orthonormal Basis.

Basis : Add how probable a signal is and subtracts energy $\Rightarrow$ if a signal has a lot of energy then we should have a