III. Hypothesis Testing (Part B)

❖ **Goal:** how to decide:

$\left\{\begin{array}{l}\text{an event has occurred}\\ \text{a signal is present}\end{array}\right.$ **?**

We need the ability to make a decision among several choices.

Basic Probability concepts

a-priori/posteriori probability

Bayes Rule

MAP detection

Bayes detection

Error types

Maximum likelihood criterion

Maximum error probability criterion

MinMax criterion

Neyman-Pearson criterion

Multiple hypotheses

Composite hypotheses testing

Receiver Operator Characteristic (ROC) curves

❖ Neyman-Pearson Criterion

- *a priori* probabilities
- cost of each type of error

$\left.\begin{array}{l}\\ \\ \end{array}\right\}$ may be difficult to determine

Radar Applications $\rightarrow$ what is important to know? probability of detection probability of false alarm

- **Goal:** determine decision rule where

  $\hookrightarrow$ we maximize the probability of detection for a given probability of false alarm.

$$\Rightarrow \max P\left(D_1 \middle| H_1\right) \quad \text{subject to} \quad P\left(D_1 \middle| H_0\right) = \alpha$$

Recall: $P\left(D_1 \middle| H_1\right) = 1 - P\left(D_0 \middle| H_1\right)$

$$\Rightarrow \min P\left(D_0 \middle| H_1\right) \quad \text{subject to} \quad P\left(D_1 \middle| H_0\right) = \alpha$$

- Use Lagrange multipliers

$$C = P\left(D_0 \middle| H_1\right) + \lambda\left[P\left(D_1 \middle| H_0\right) - \alpha\right]$$

called Lagrange multiplier

$$C = P\left(D_0 \middle| H_1\right) + \lambda\left[P\left(D_1 \middle| H_0\right) - \alpha\right]$$

Example:

$$H_0: \quad y = n \qquad\qquad n \sim N(0,1)$$
$$H_1: \quad y = s + n \qquad\qquad s = 1$$

Design the optimum decision rule with a fixed $P_{FA} = 0.1$ based on one sample

$$f_1(y) =$$

$$f_0(y) =$$

$$\Lambda(y) =$$

Example:

$$H_0: \quad y(n) = w(n) \qquad n = 0, \cdots, 99$$
$$H_1: \quad y(n) = s + w(n) \qquad n = 0, \cdots, 99$$
$$w \sim N(0,1)$$

Design the optimum decision rule with a fixed $P_{FA} = 0.1$ (based on 100 samples), s=1.

$$f_1(y) =$$

$$f_0(y) =$$

$$\Lambda(y) =$$

❖ Multiple Hypotheses Testing

- We may have more than two decisions possible.

- Assume – cost function C_{ij} is known
  – *a priori* probabilities are known

- Cost for hypothesis H_j given sample y is:

  $C_j =$

- Average cost is:

  $C =$

- **Example:**

Assume three possible signals may be present.

$$H_0: \quad y_n \sim N\left(m_0, \sigma^2\right)$$

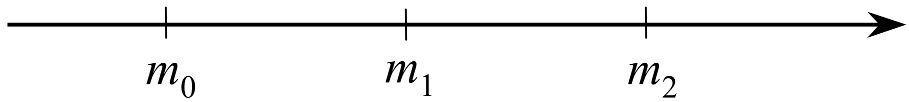
$$H_1: \quad y_n \sim N\left(m_1, \sigma^2\right)$$

$$H_2: \quad y_n \sim N\left(m_2, \sigma^2\right)$$

Assume (1) $P_i = 1/3$, $C_{ii} = 0$, $C_{ij} = 1 \quad i \neq j$

(2) N samples are available

Design the decision rule which minimizes the probability of error using *1* sample.


m_0
m_1
m_2

❖ Composite Hypothesis Testing

- We may not know the signal precisely (phase, amplitude).
- Uncertainty in parameters may be built in the detection scheme.

<u>Example</u>: H_0 : no signal

H_1 : signal with phase θ_1

H_2 : signal with phase θ_2, etc.

Note:
- detection $\Rightarrow$ estimation of parameters
- estimation may not be needed
- $\rightarrow$ need to average "uncertainties" out

- Approach:
  - Model unknown parameter(s) as a random variable.
  - Compute decision rule as a function of unknown parameter.
  - Average decision rule with respect to unknown parameter(s).

- **Example:**

  Assume a binary detection where

$$
\begin{aligned}
H_0 &: \quad y_n = w_n \\
H_1 &: \quad y_n = m + w_n
\end{aligned}
\qquad\qquad w_n \sim N\left(0, \sigma^2\right)
$$

  Assume m is a RV with Normal pdf $N(0,\sigma^2_m)$.

$f_m(m) =$

  Compute the decision rule.

❖ What if the random parameter has an unknown pdf ?

→ Estimate it or take a worse case scenario, and follow as in example before

❖ What if the unknown parameter is deterministic ?

- No pdf can be used
- What to do:

° Use a Neyman Pearson test: see whether the test decision rule depends on the unknown parameter

⇒ <u>if it does not</u>: we're done !
(test called *Uniformly Most Powerful* (*UMP*) test)

⇒ <u>if it does</u>: estimate unknown parameter first

° estimated parameter used in a "*generalized LRT*"

- **Example:**

Assume we have

H_0: $y \sim N(0,1)$

H_1: $y \sim N(m,1)$

Assume *m* is a deterministic unknown positive value

Compute the decision rule based on one sample, assume a given P_{FA}

How does P_D vary as a function of *m* ?

- **Example:**

Assume we have

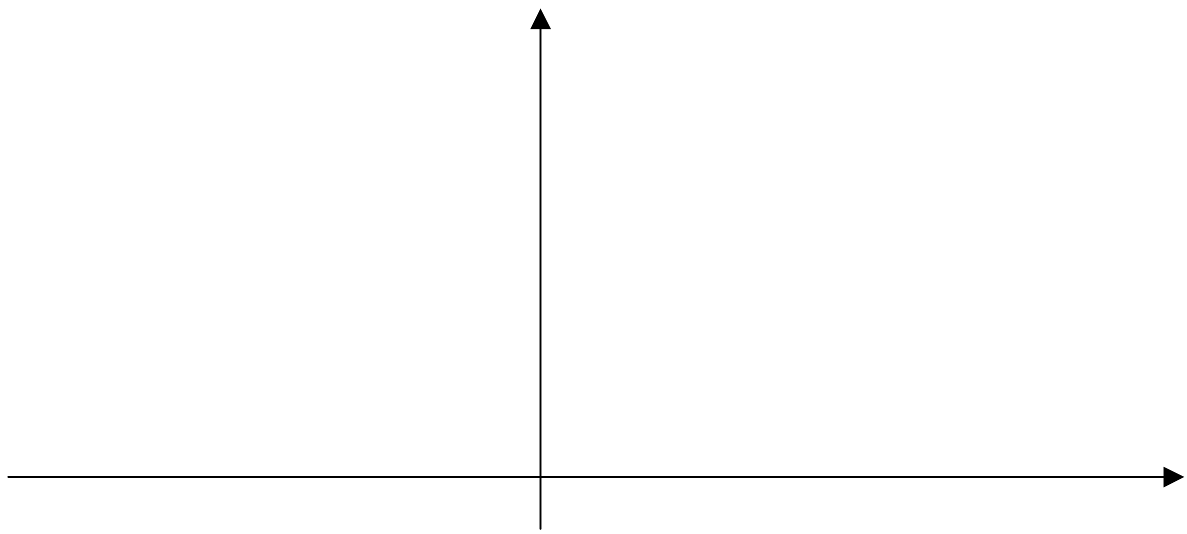
H_0: $y \sim N(0,1)$

H_1: $y \sim N(m,1)$

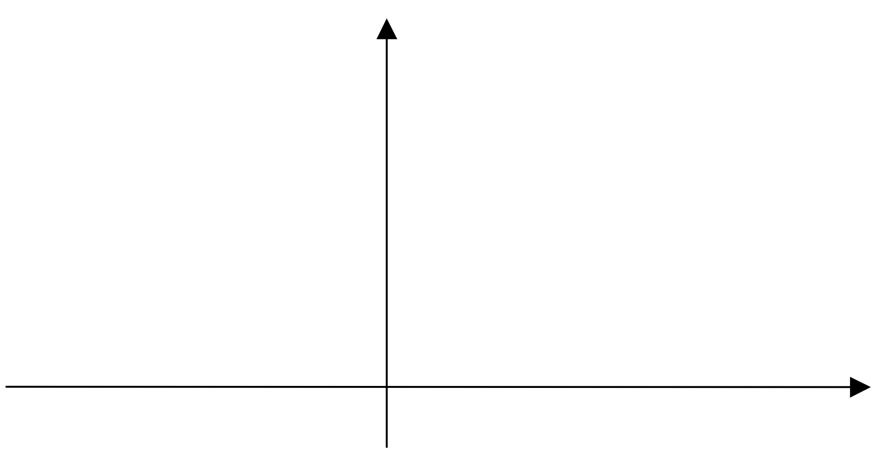
Assume *m* is a deterministic unknown value (can be either positive or negative)

Compute the decision rule based on one sample, assume a given P_{FA}

• **Detection performance as a function of *m***



- **Alternative: test based on |*m*|**



- **Example:**

Assume we have

H_0: $y_n \sim N(0,1)$

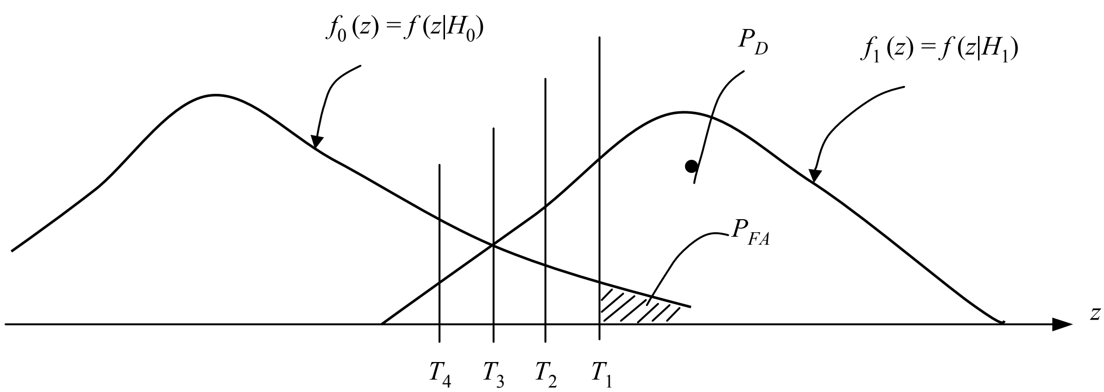
H_1: $y_n \sim N(m,1),\ n=1,\ \ldots,\ N$

Assume m is a deterministic unknown positive value

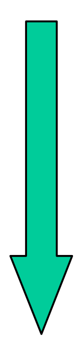
Compute the decision rule, based on N samples

❖ Receiver Operator Characteristic (ROC) Curves

- ROC curves are used to evaluate the performance of the receiver.
- ROC curves given in terms of P_{FA} and P_s.



- What does the ROC curve look like ?



- **Look at the earlier example and derive expression for P_D and P_{FA}**

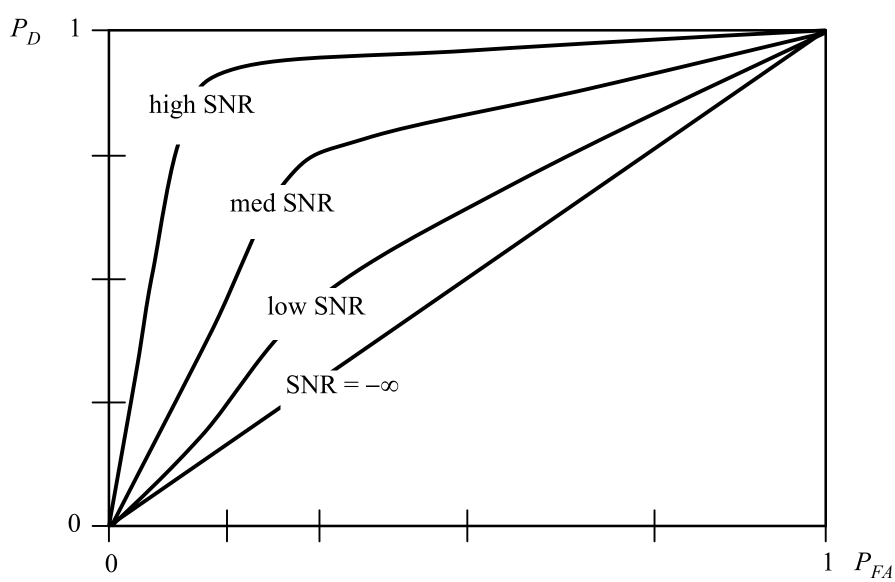
- **Use above results to derive general trend for ROC curves**

  Assume we have

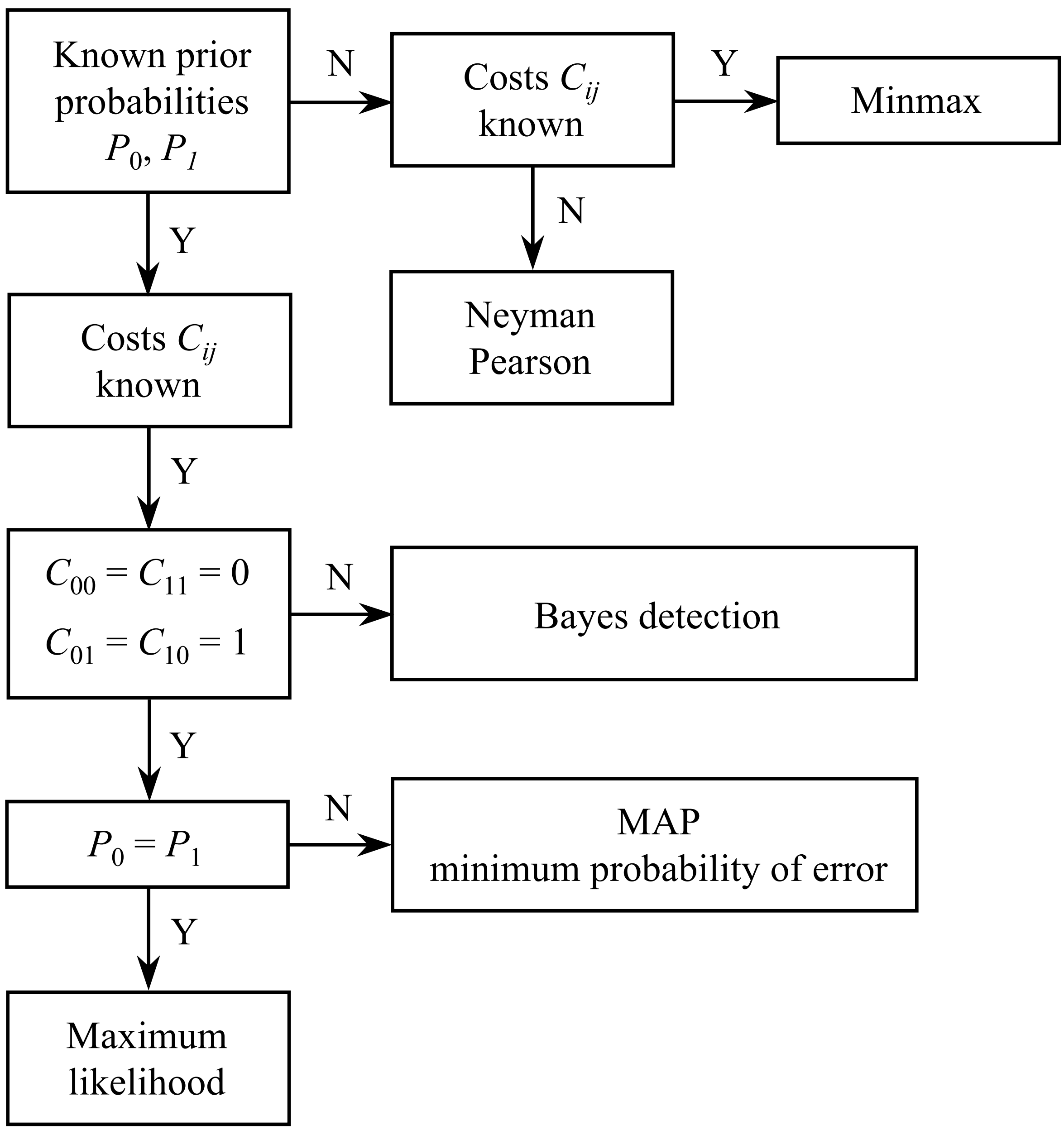
  H_0: $y \sim N(0,\sigma^2)$

  H_1: $y \sim N(m,\sigma^2)$

  Assume m is a deterministic unknown positive value


P_D
1
high SNR
med SNR
low SNR
SNR = –∞
0
0
1
P_{FA}

Binary Hypothesis Testing Schemes Summary



Test Name	Data Model Assumptions	Decision Rule	Optimality Criterion
Minimum probability of error (MAP)	• Hypothesis modeled as random events with known pdfs. • Known prior probabilities P_0, P_1.	$\dfrac{f_1(y)}{f_0(y)} \underset{H_0}{\overset{H_1}{\gtrless}} \dfrac{P_0}{P_1}$	Minimize cost $C = P_0 P(D_1 \vert H_0) + P_1 P(D_0 \vert H_0)$
Maximum likelihood	• Hypothesis modeled as random events with known pdfs.	$\dfrac{f_1(y)}{f_0(y)} \underset{H_0}{\overset{H_1}{\gtrless}} 1$	
Bayes detection	• Hypothesis modeled as random events with known pdfs. • Known prior probabilities P_0, P_1. • Cost functions C_{ij} known.	$\dfrac{f_1(y)}{f_0(y)} \underset{H_0}{\overset{H_1}{\gtrless}} \dfrac{P_0(C_{10}-C_{00})}{P_1(C_{01}-C_{11})}$	Minimize cost $C = \sum_{i,j=0}^{1} C_{ij} P(D_i \vert H_j) P(H_j)$
Minmax	• Hypothesis modeled as random events with known pdfs. • Cost functions C_{ij} known.	$\dfrac{f_1(y)}{f_0(y)} \underset{H_0}{\overset{H_1}{\gtrless}} \gamma$ where γ defined so that $P_{FA} = \dfrac{C_{11}-C_{00}}{C_{10}-C_{00}} + \dfrac{C_{01}-C_{11}}{C_{10}-C_{00}} P_M$ with $P_{FA} = \int_{\gamma}^{\infty} f_0(y)\,dy$ $P_M = \int_{-\infty}^{\gamma} f_1(y)\,dy$	Minimize maximum average cost $\dfrac{\partial P}{\partial P_1} = 0 ==>$ $(C_{11}-C_{00}) + (C_{01}-C_{11}) P_M$ $-(C_{10}-C_{00}) P_{FA} = 0$
Neyman-Pearson	• Hypothesis modeled as random events with known pdfs.	$\dfrac{f_1(y)}{f_0(y)} \underset{H_0}{\overset{H_1}{\gtrless}} \gamma$ with $P_{FA} = \alpha$ user specified	Maximize probability of detection P_D for a given P_{FA}