

# ET4386 Estimation and Detection

Detection

Lecture1: Introduction Detection Theory (Ch 1,2,3)

**Dr. Justin Dauwels** ([j.h.g.dauwels@tudelft.nl](mailto:j.h.g.dauwels@tudelft.nl))

# Learning Objectives

- LO1: Understand the definition of **detection**
- LO2: Get familiar with important **probability density functions (pdfs)**
  - Gaussian pdf
  - Central Chi-squared pdf
- LO3: Get familiar with **optimal binary detection**:
  - Neyman-Pearson Theorem
  - Minimum Probability of Error
  - Bayesian Risk

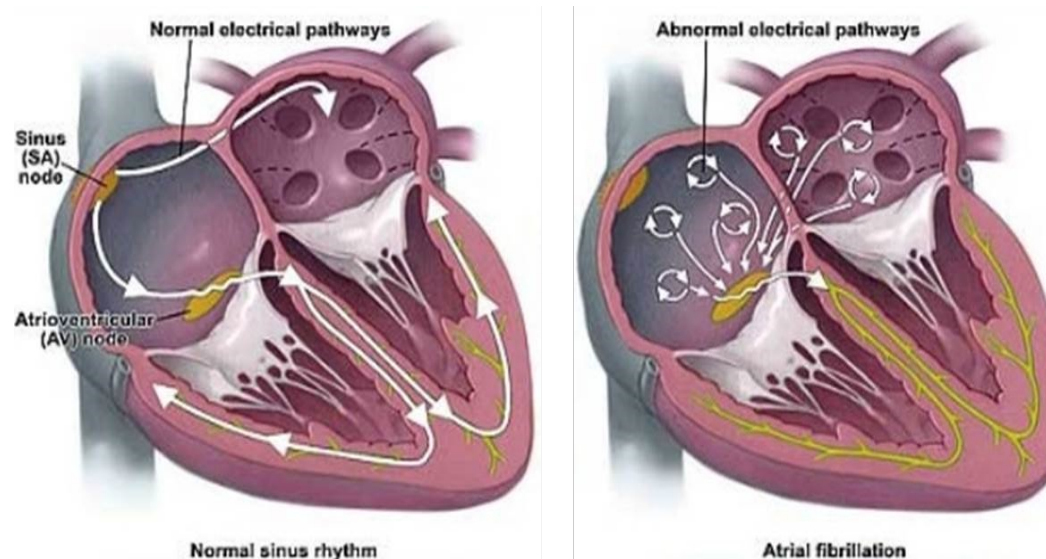
# Overview

- Introduction to **detection theory**
- Important **probability density functions (pdfs)**
  - Gaussian pdf
  - Central Chi-squared pdf
- Basics of **optimal binary detection**:
  - Neyman-Pearson Theorem (today)
  - Minimum Probability of Error (today)
  - Bayesian Risk (next time)

# Biomedical Signal Processing

## – Detection of cell activity in Atrial Fibrillation

- Atrial fibrillation (AF) is one of the most common age related cardiac arrhythmia
- AF is characterized by rapid and irregular electrical activity of the atria.
- AF is rooted in impaired electrical conduction known as electropathology.



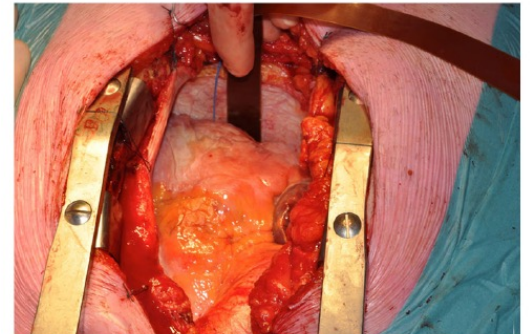
# Biomedical Signal Processing

## – Detection of cell activity in Atrial Fibrillation

Common treatment: Ablation. Only successful in 50 % of cases...

Research project on atrial fibrillation:

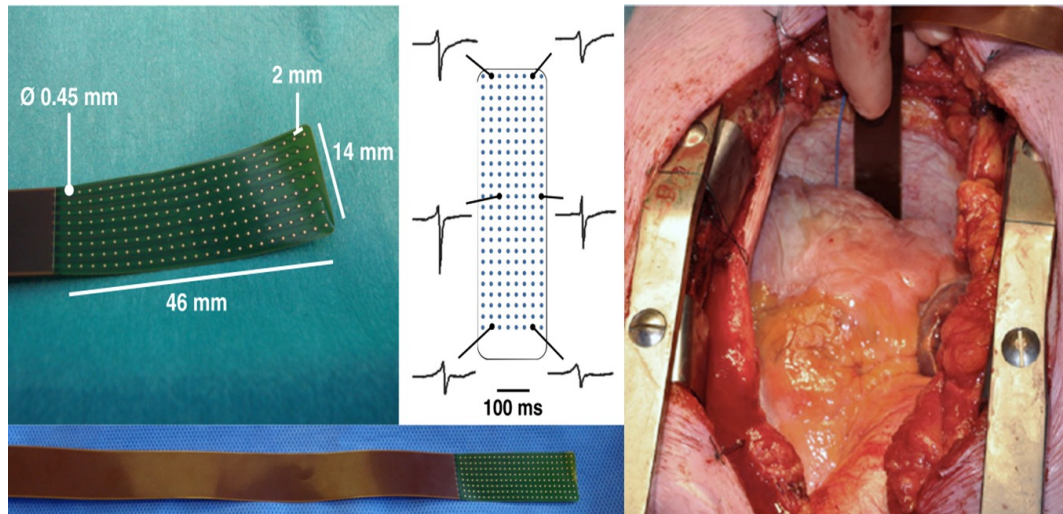
- Measure signals directly on the heart using a high dimensional sensor (188).
- Get understanding of the real problem in Atrial fibrillation.
- Develop less invasive methods.



# Biomedical Signal Processing

## – Detection of cell activity in Atrial Fibrillation

An electrogram is a record of changes in the electrical potential of the (many) cells in the neighborhood of an electrode that is positioned on the heart surface



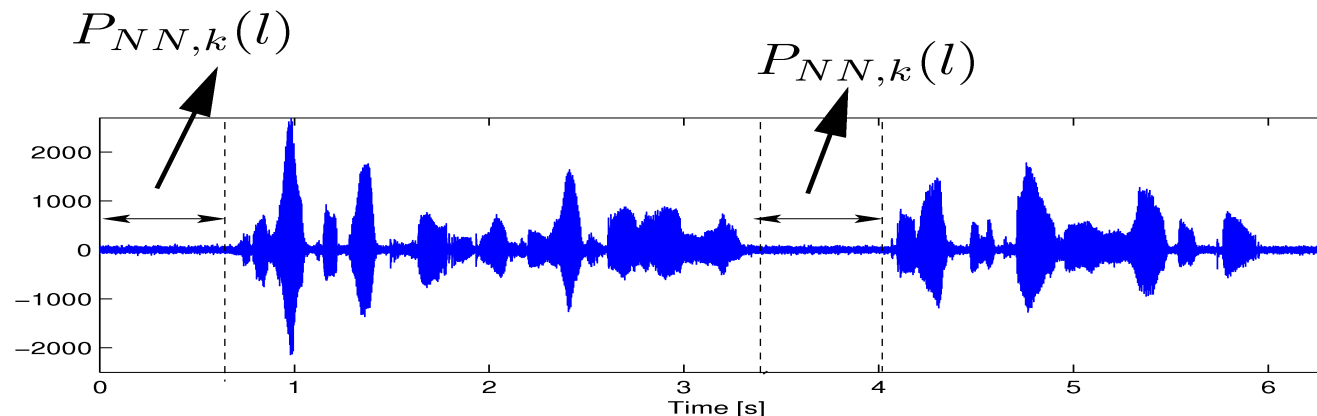
To understand which (group of) cells are problematic, we would like to *detect* when each cell is being activated and *estimate* the underlying conductivity of the cell.

# Example 2 – Speech Processing

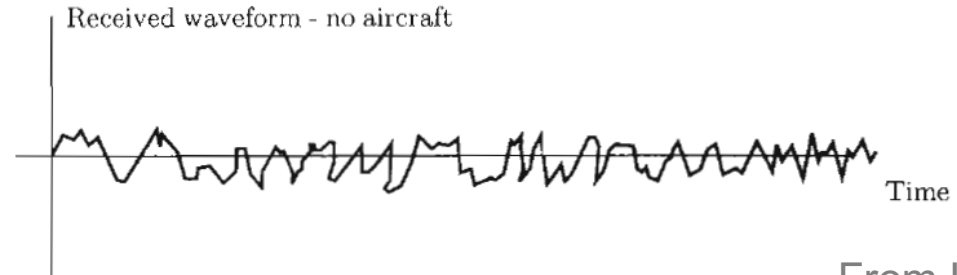
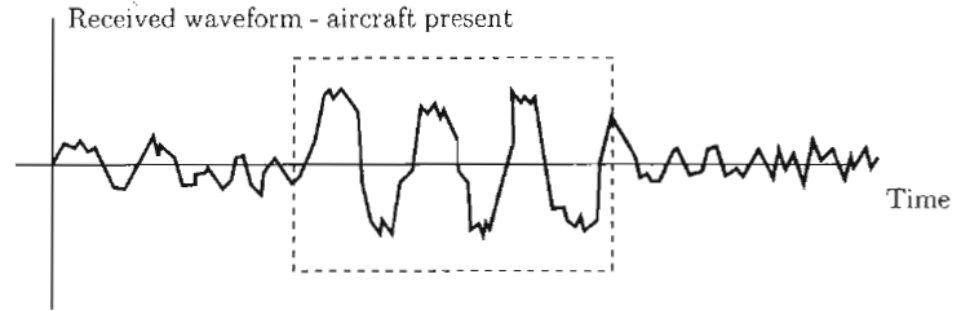
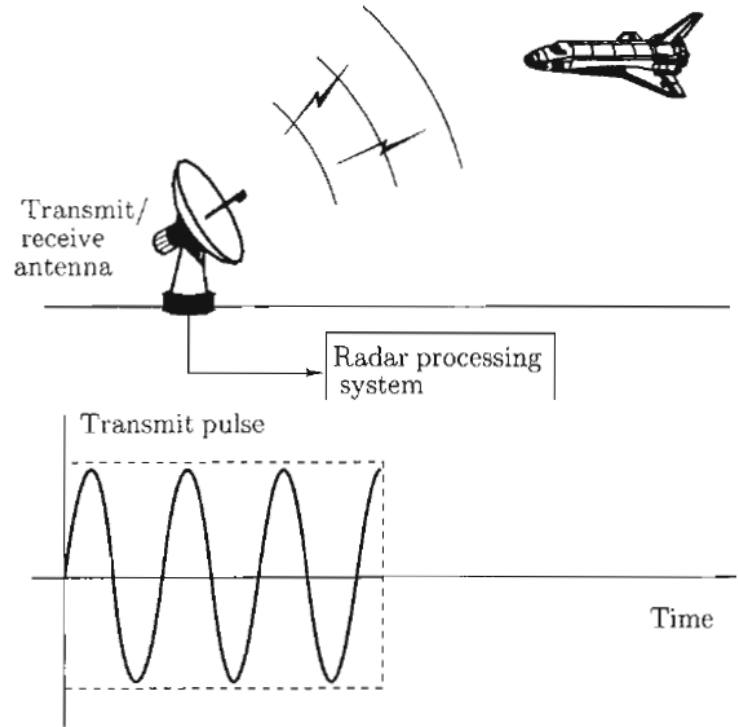
Voice activity detection (VAD):

In speech processing applications a VAD is commonly used, e.g.,

- In speech enhancement: to determine whether speech is present or not. If speech is not present, the remaining signal consists of noise only and can be used to estimate the noise statistics.
- Speech coding: Detect whether speech is present. If speech is not present, there is no need for the device (phone) to transmit any information.



# Example 3 – Radar



From Kay, part II.

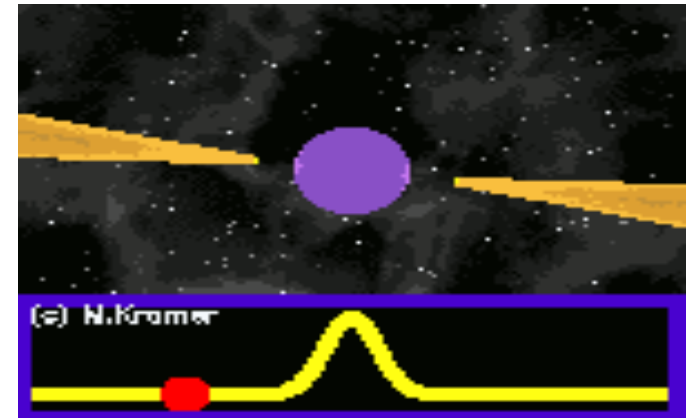
- Detecting the presence of an aircraft using Radar.
- Transmit a known signal, analyse received waveform.



# Example 4 – Radio Pulsar Navigation

Pulsars (pulsating star):

- Highly magnetized rotating neutron star that emits a beam of electromagnetic radiation.
- Radiation can only be observed when the beam of emission is pointing toward the earth (lighthouse model)
- wideband (100 Mhz - 85 Ghz)
- extremely accurate pulse sources.



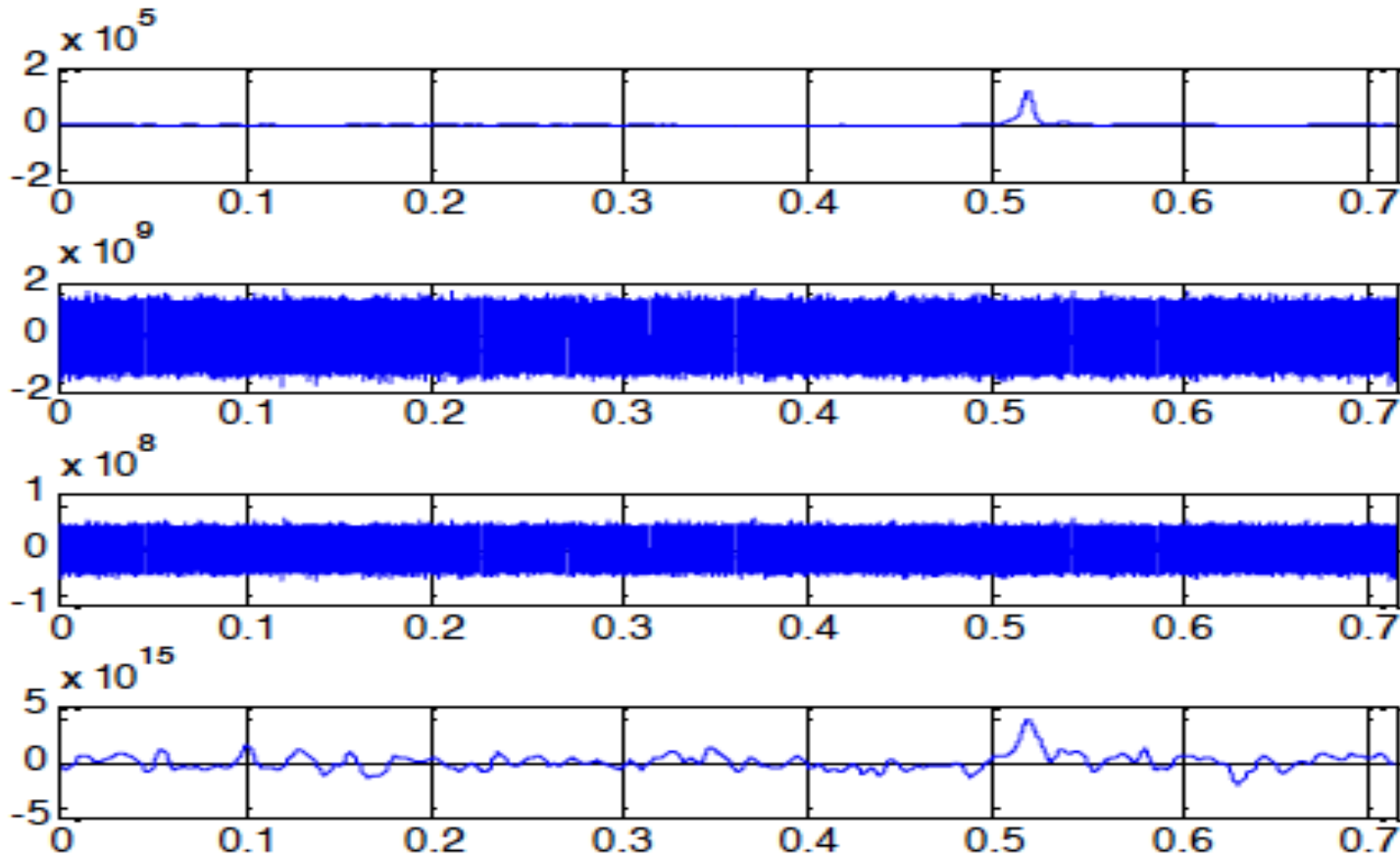
*Kramer (University of Manchester)*

# Example 4 – Radio Pulsar Navigation

For some millisecond pulsars, the regularity of pulsation is more precise than an atomic clock.

- Pulsars are "ideal" for time-of-arrival
- pulsar signals are weak (SNR = -90 dB)

# Example 4 – Radio Pulsar Navigation



pulsar B0329+54

observation (SNR = -90 dB)

epoch folding ( $K = 10^3$ )

matched filtering

# What is Detection Theory?

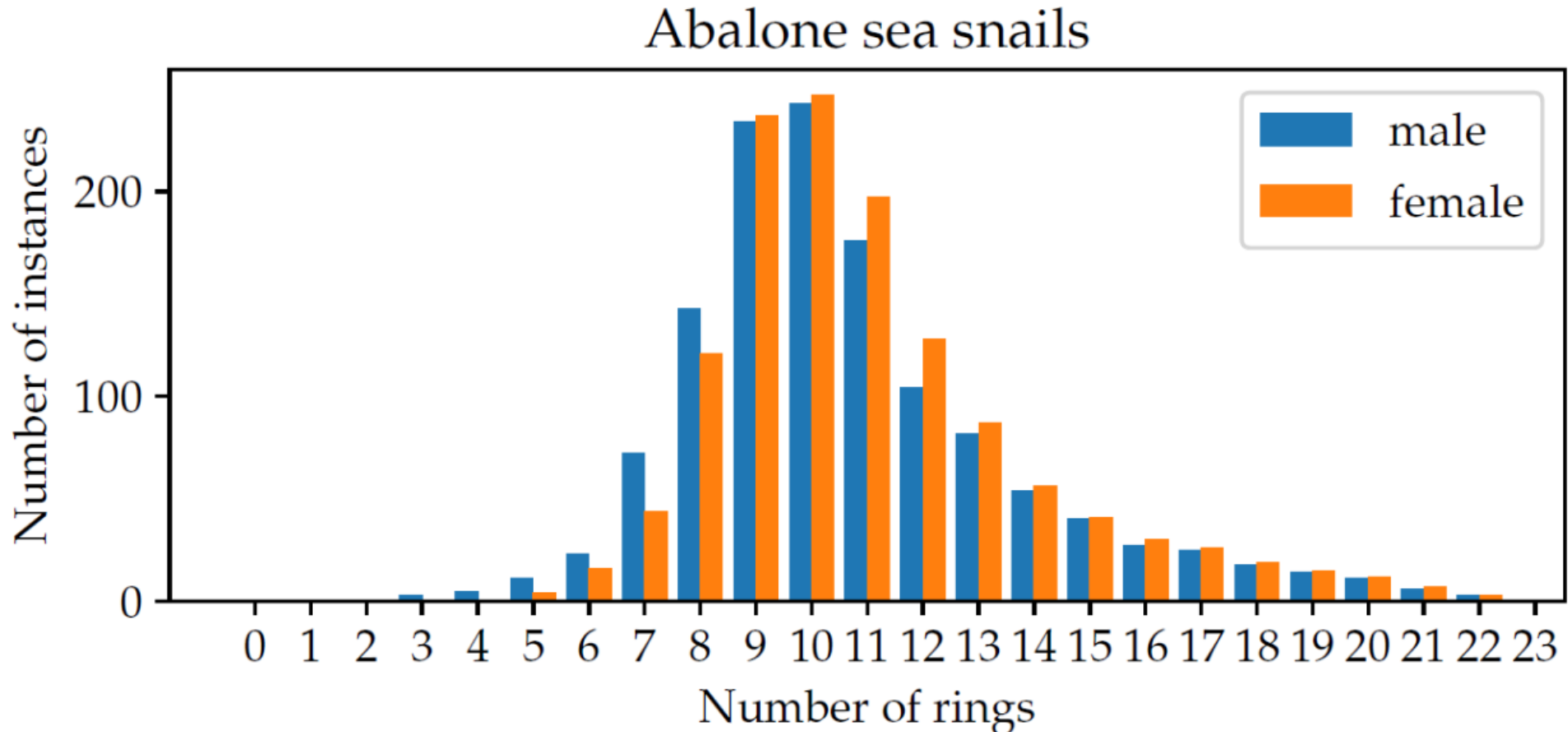
Abalone



Is it a male or female abalone?

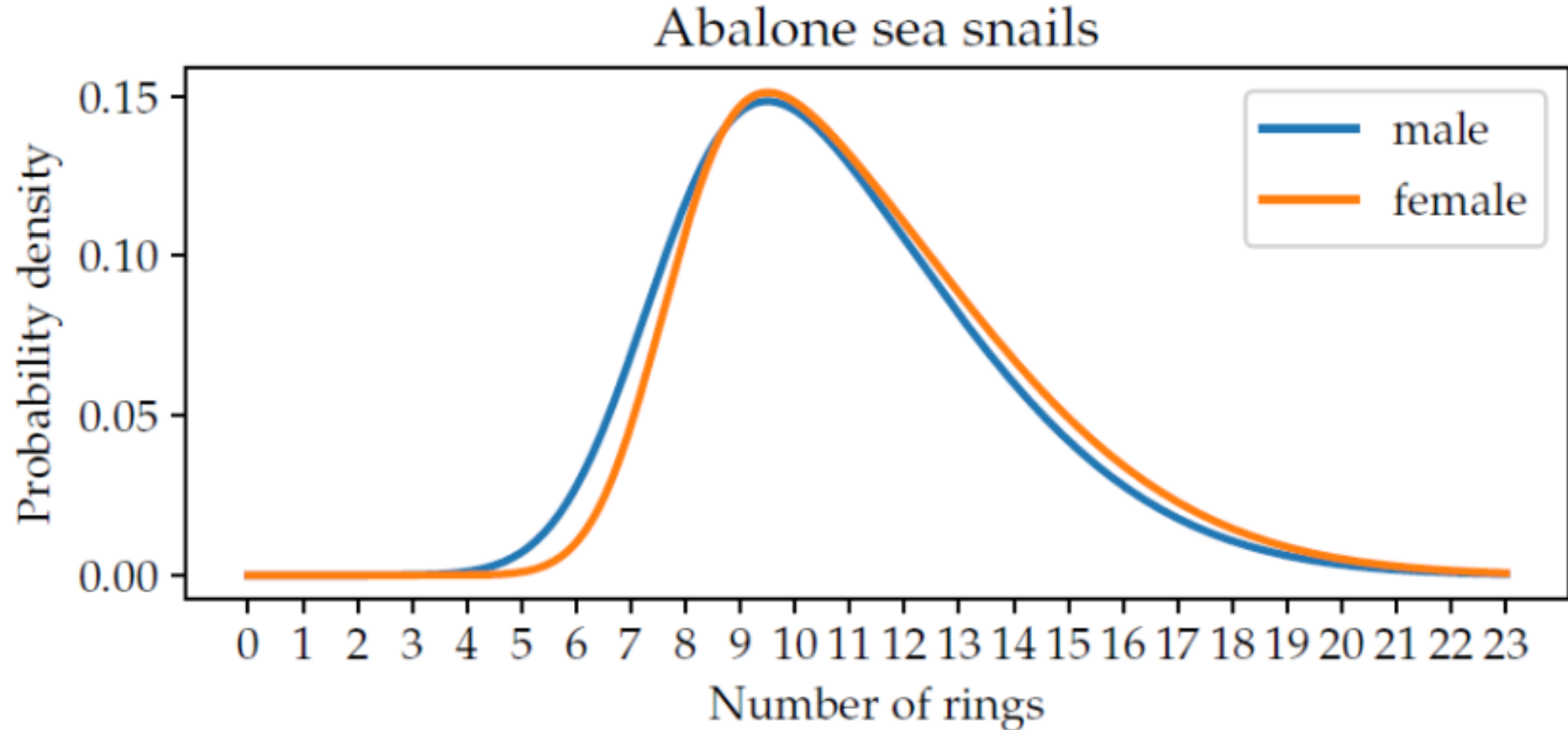
# What is Detection Theory? (2)

Is it a male or female abalone?



# What is Detection Theory? (3)

Is it a male or female abalone?



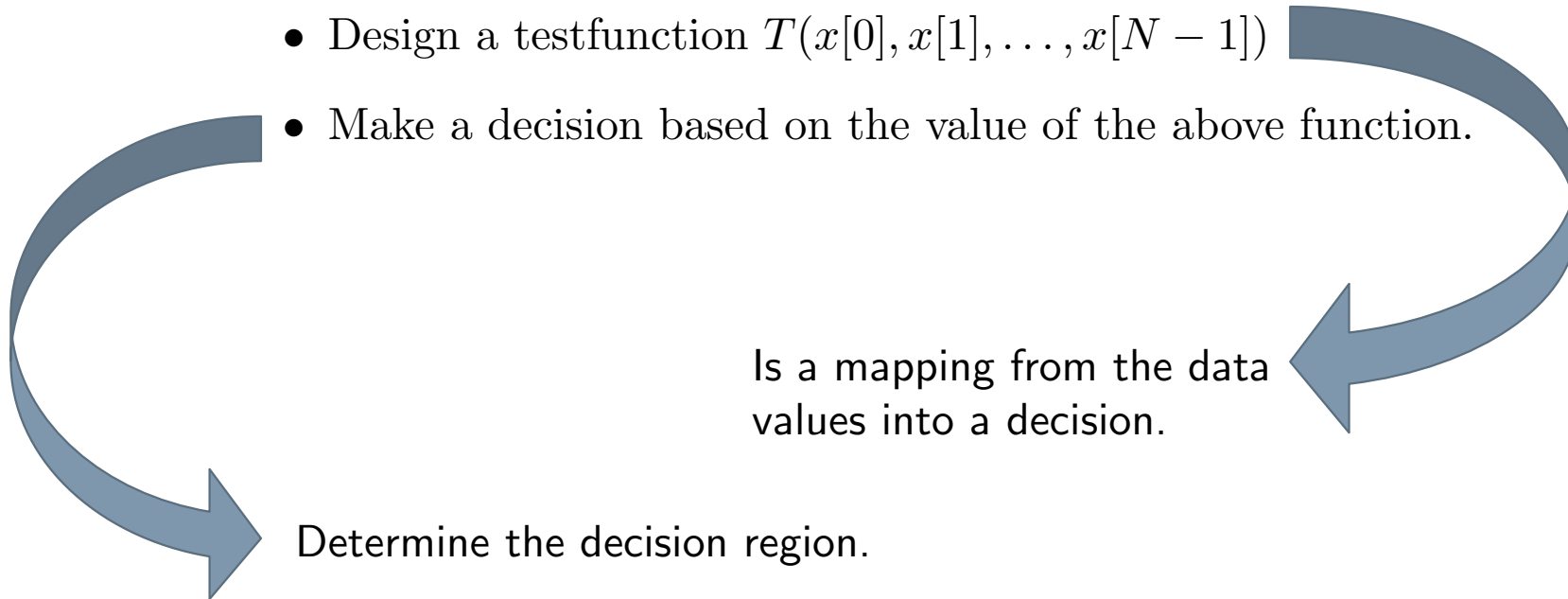
# What is Detection Theory? (3)

Given a set of measurements  $x[0], x[1], \dots, x[N - 1]$

- Design a testfunction  $T(x[0], x[1], \dots, x[N - 1])$
- Make a decision based on the value of the above function.

Is a mapping from the data values into a decision.

Determine the decision region.



# The Simplest Detection Problem

Binary detection: Determine whether a certain signal that is embedded in noise is present or not.

$$\begin{aligned}\mathcal{H}_0 \quad x[n] &= w[n] \\ \mathcal{H}_1 \quad x[n] &= s[n] + w[n]\end{aligned}$$

Note that if the number of hypotheses is more than two, then the problem becomes a multiple hypothesis testing problem. One example is detection of different digits in speech processing.

- $x[n]$  is a single sample measurement
- $s[n]$  is the signal of interest and  $w[n]$  the noise, e.g,  $w[n] \sim N(0, \sigma^2)$
- $\mathcal{H}_0$  (signal absent) is the Null hypothesis
- $\mathcal{H}_1$  (signal present) is the Alternative hypothesis
- we assume distribution of  $p(x; \mathcal{H}_0)$  and  $p(x; \mathcal{H}_1)$  are known.



# Example

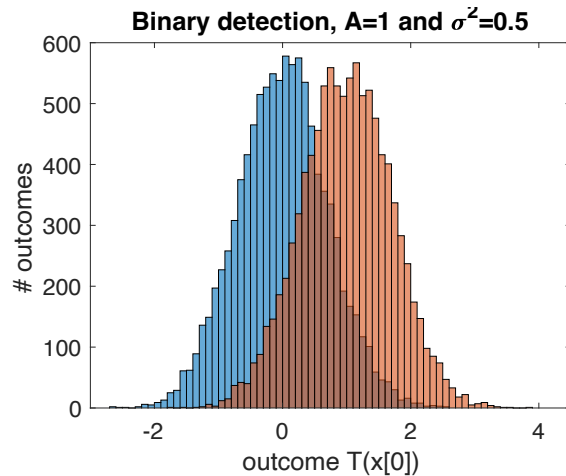
Detection of a DC level of amplitude  $A = 1$  embedded in white Gaussian noise  $w[n]$  with variance  $\sigma^2$  with only one sample.

$$\mathcal{H}_0 \quad : \quad x[0] = w[0]$$

$$\mathcal{H}_1 \quad : \quad x[0] = 1 + w[0]$$

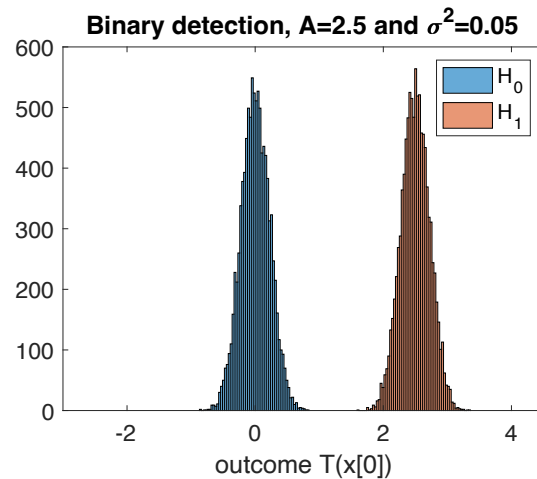
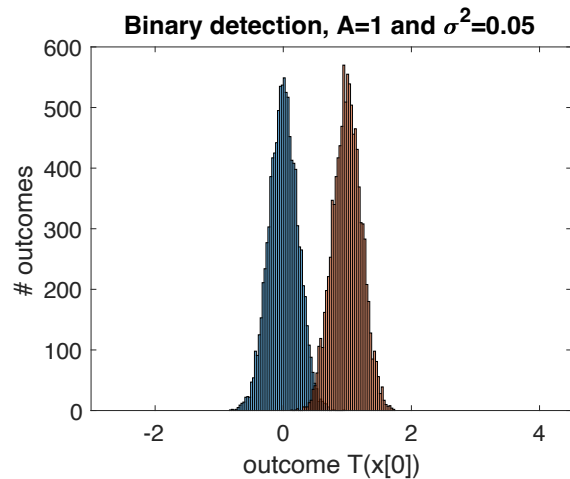
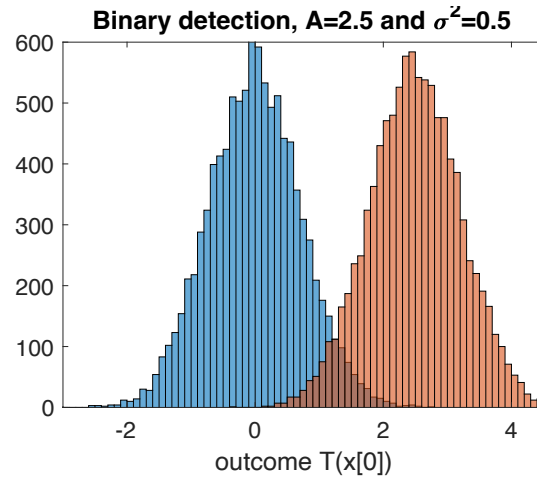
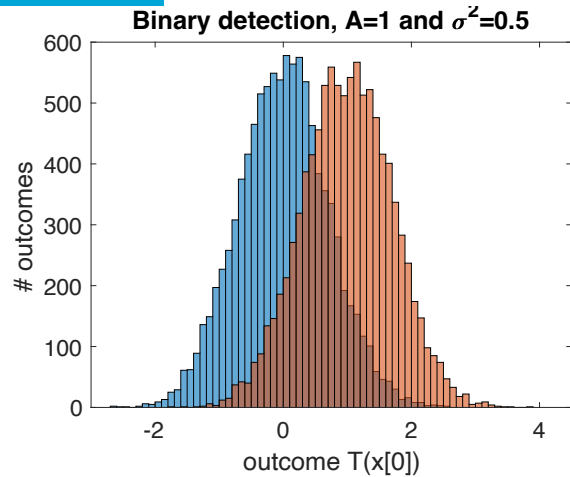
Imagine that  $x[0]$  is the outcome of a test, e.g., for Covid-19. How to decide whether someone has COVID-19 ( $\mathcal{H}_1$ ) or not ( $\mathcal{H}_0$ )?

# Example



- What determines how well we can distinguish  $\mathcal{H}_0$  and  $\mathcal{H}_1$ ?
- What would be a good threshold?
- What would happen if we increase or decrease this threshold?
- What is the implication of a wrong decision?
  - What is the cost of a miss?
  - A false alarm?

# Example



- What determines how well we can distinguish  $\mathcal{H}_0$  and  $\mathcal{H}_1$ ?
- What would be a good threshold?
- What would happen if we increase or decrease this threshold?
- What is the implication of a wrong decision?
  - What is the cost of a miss?
  - A false alarm?

# Example

Detection of a DC level of amplitude  $A = 1$  embedded in white Gaussian noise  $w[n]$  with variance  $\sigma^2$  with  $N$  samples:

$$\mathcal{H}_0 : x[n] = w[n] \quad n = 0, 1, \dots, N - 1$$

$$\mathcal{H}_1 : x[n] = 1 + w[n] \quad n = 0, 1, \dots, N - 1$$

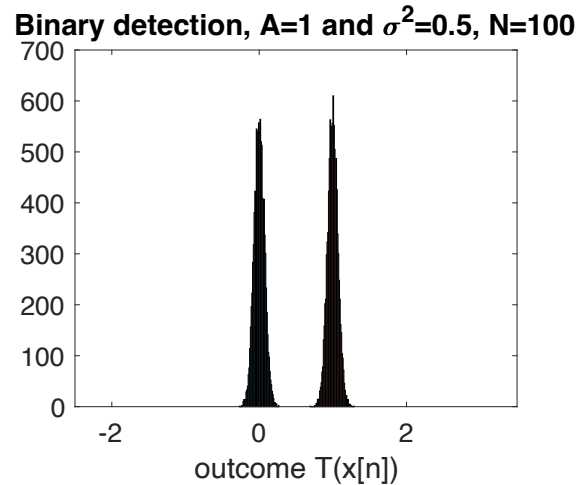
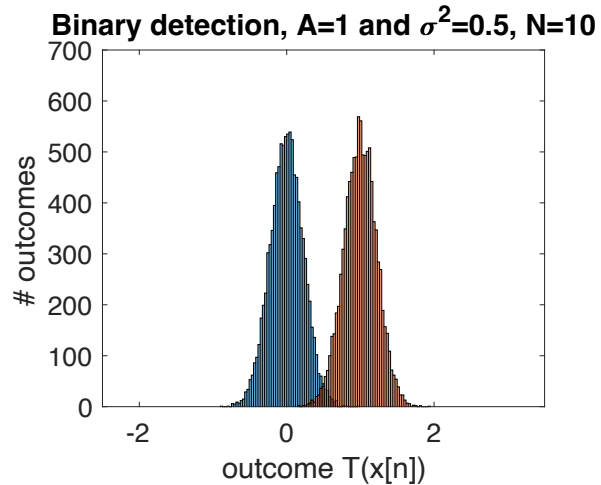
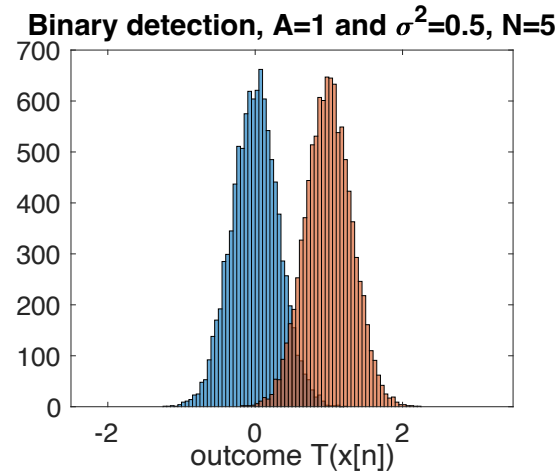
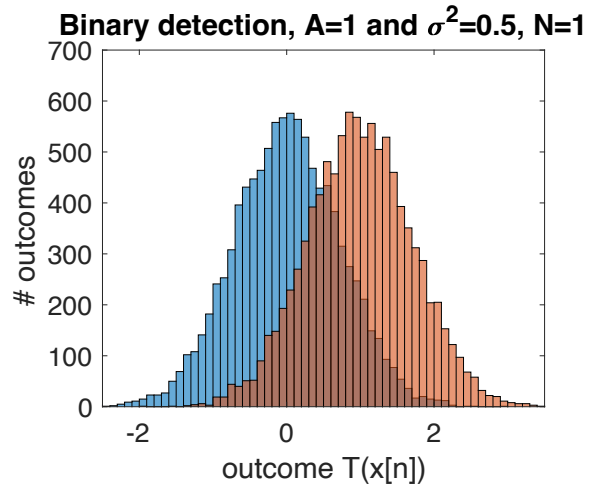
How should we form  $T(x[n])$ ?

A possible detector for choosing  $\mathcal{H}_1$  is

$$T(x[n]) = \frac{1}{N} \sum_{n=0}^{N-1} x[n] \geq \gamma.$$

- Distribution of  $T(x[n])$ ?
- effect of # of samples?

# Example



$$T(x[n]) = \frac{1}{N} \sum_{n=0}^{N-1} x[n] \geq \gamma.$$

- Distribution of  $T(x[n])$ ?
- effect of # of samples?



0/0

Question sl

## Q2: The increase of which parameters lead to better detection? (Multiple choice)

**A**

0%

**$\sigma$**

0%

**N**

0%

Join at: [vevox.app](https://vevox.app)

ID: 135-988-786



0/0

Preparing Re

## Q2: The increase of which parameters lead to better detection? (Multiple choice)

**A**



0%

**$\sigma$**

0%

**N**



0%

Join at: [vevox.app](https://vevox.app)

ID: 135-988-786

# Detection Performance

- Can we expect to always make a correct decision? Depending on the noise variance  $\sigma^2$ , it will be more or less likely to make a decision error.
- How to make an optimal decision?
- The detection performance improves with "distance" between the pdfs under  $\mathcal{H}_0$  and  $\mathcal{H}_1$ :
  - Increasing the SNR:  $\frac{A^2}{\sigma^2}$ .
  - Increasing  $N$ .
- Performance measure (for uncorrelated Gaussian data): deflection coefficient

$$d^2 = \frac{(E(T; \mathcal{H}_1) - E(T; \mathcal{H}_0))^2}{\text{var}(T; \mathcal{H}_0)} = \frac{NA^2}{\sigma^2}$$



# Overview

- Introduction to **detection theory**
- Important **probability density functions (pdfs)**
  - Gaussian pdf
  - Central Chi-squared pdf
- Basics of **optimal binary detection**:
  - Neyman-Pearson Theorem (today)
  - Minimum Probability of Error (next time)
  - Bayesian Risk (next time)

# Important pdfs – Gaussian pdf

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x - \mu)^2\right) \quad -\infty < x < +\infty$$

where  $\mu$  is the mean and  $\sigma^2$  is the variance of  $x$ .

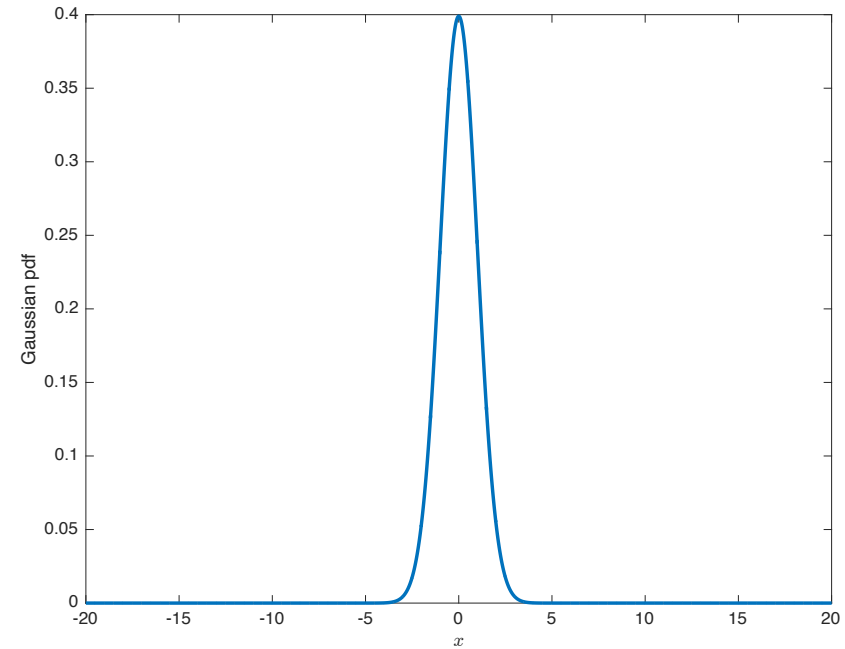
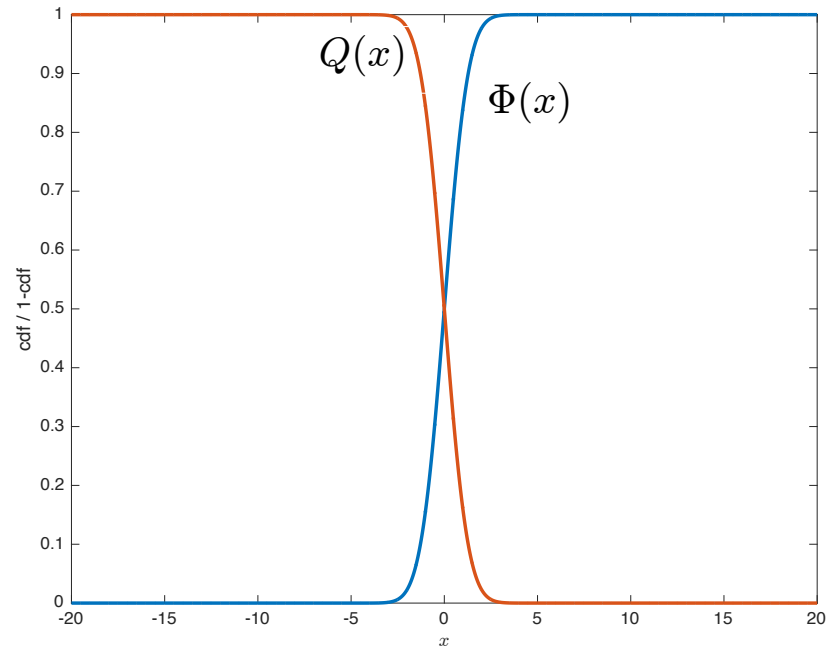
Standard normal pdf:  $\mu = 0$  and  $\sigma^2 = 1$

The cumulative distribution function (cdf) of a standard normal pdf:

$$\Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}t^2\right) dt$$

A more convenient description is the right-tail probability which is defined as  $Q(x) = 1 - \Phi(x)$ . This function which is called **Q-function** is used frequently in different detection problems where the signal and noise are normally distributed.

# Important pdfs – Gaussian pdf



# Using the Q-function for the Gaussian pdf

- For  $T(x[n]) = \frac{1}{N} \sum_{n=0}^{N-1} x[n]$ , evaluate  $P(T(x[n]) \geq \gamma)$  when  $x \sim \mathcal{N}(\mu, \sigma^2)$ .
- $P(T(x[n]) \geq \gamma) = P(\frac{1}{N} \sum_{n=0}^{N-1} x[n] \geq \gamma) = Q\left(\frac{\gamma - \mu}{\sqrt{\sigma^2/N}}\right)$ , where

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

is the right-tail probability of the Gaussian PDF.

# Important pdfs – central Chi-squared

A chi-squared pdf arises as the pdf of  $x$ , where  $x = \sum_{i=1}^v x_i^2$ , if  $x_i$  is a standard normally distributed random variable. The chi-squared pdf with  $v$  degrees of freedom is defined as

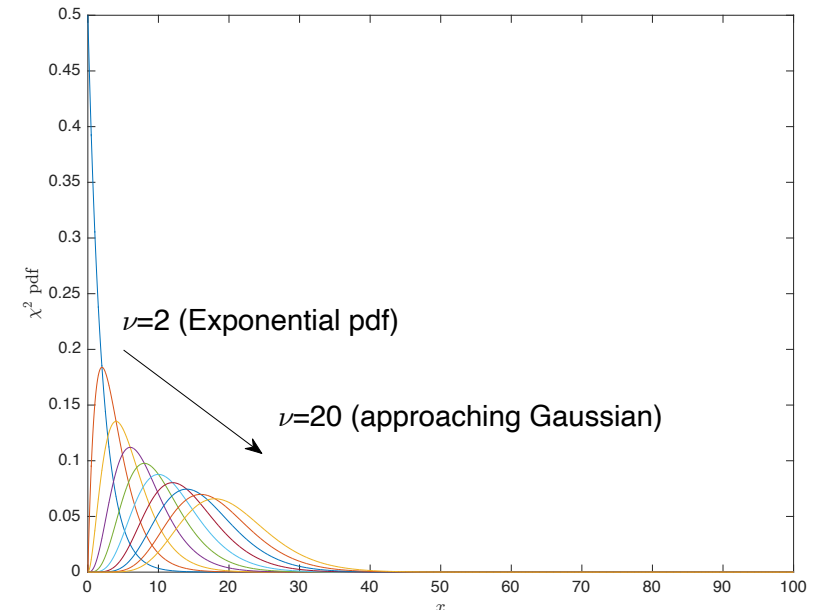
$$p(x) = \begin{cases} \frac{1}{2^{\frac{v}{2}} \Gamma(\frac{v}{2})} x^{\frac{v}{2}-1} \exp\left(-\frac{1}{2}x\right), & x > 0 \\ 0, & x < 0 \end{cases}$$

and is denoted by  $\chi_v^2$ .  $v$  is assumed to be integer and  $v \geq 1$ . The function  $\Gamma(u)$  is the Gamma function and is defined as

$$\Gamma(u) = \int_0^{\infty} t^{u-1} \exp(-t) dt$$

The right-tail probability for a  $\chi_v^2$  random variable is defined as

$$Q_{\chi_v^2}(x) = \int_x^{\infty} p(t) dt.$$



# Overview

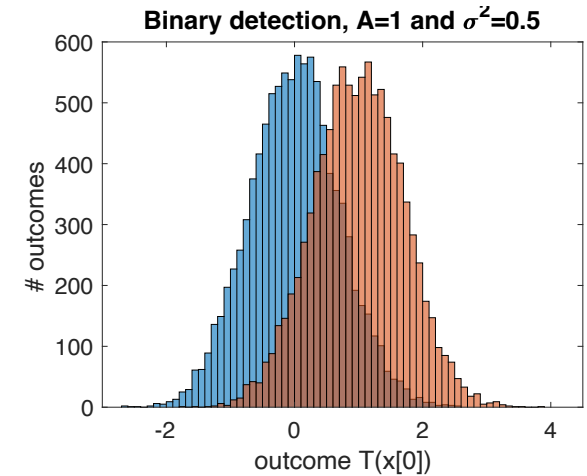
- Introduction to **detection theory**
- Important **probability density functions (pdfs)**
  - Gaussian pdf
  - Central Chi-squared pdf
- Basics of **optimal binary detection**:
  - Neyman-Pearson Theorem (today)
  - Minimum Probability of Error (today)
  - Bayesian Risk (next time)

# Making Optimal Decisions

Remember the example:

$$\mathcal{H}_0 : x[0] < \gamma$$

$$\mathcal{H}_1 : x[0] > \gamma$$



Using detection theory, rules can be derived on how to chose  $\gamma$ .

- Neyman-Pearson Theorem: Maximize detection probability for a given false alarm probability.
- Minimum probability of error
- Bayesian detector

# Neyman-Pearson Theorem - Introduction

- Simple binary detection:  $A = 1$  and  $\sigma^2 = 1$ ,  $N = 1$ .
- Under  $\mathcal{H}_0$ ,  $x[0] \sim \mathcal{N}(0, 1)$  and under  $\mathcal{H}_1$   $x[0] \sim \mathcal{N}(1, 1)$ .
- Top example:  $\gamma = 0.5$ , bottom example:  $\gamma = 1.5$
- Binary hypothesis:

$$\mathcal{H}_0 : \mu = 0$$

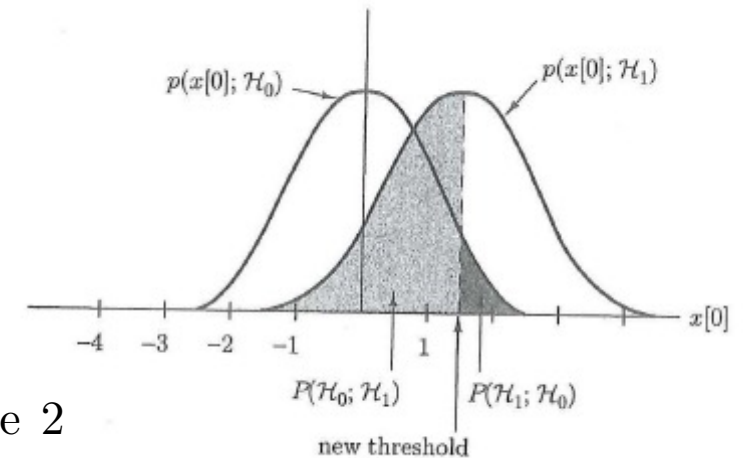
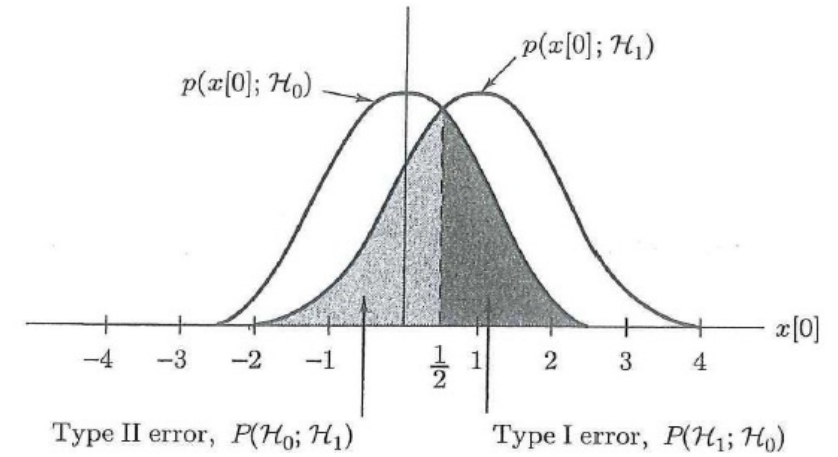
$$\mathcal{H}_1 : \mu = 1$$

Detection rule:

$$\mathcal{H}_0 : x[0] < \gamma$$

$$\mathcal{H}_1 : x[0] > \gamma$$

- Notice two type of errors: Type 1  $P(\mathcal{H}_1; \mathcal{H}_0)$  (false alarm) and Type 2  $P(\mathcal{H}_0; \mathcal{H}_1)$  (miss).



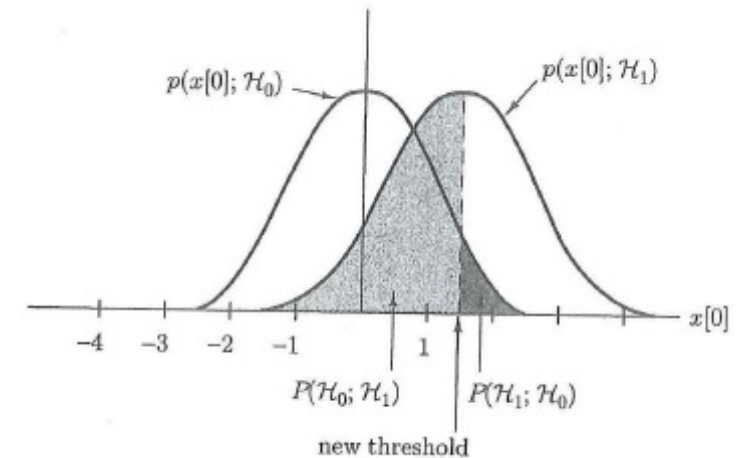
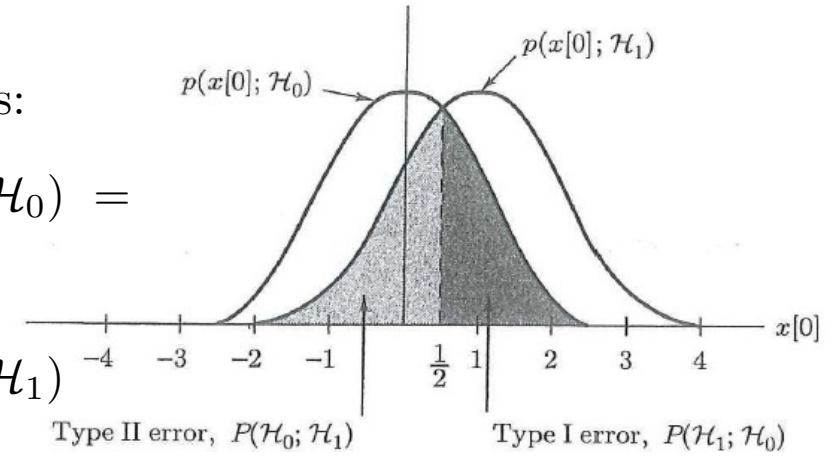
S. Kay – detection theory Figs. 3.2 and 3.3.



# Neyman-Pearson Theorem – Detection Performance

Detection performance of a system is measured mainly by two factors:

1. Probability of false alarm:  $P_{FA} = P(\mathcal{H}_1; \mathcal{H}_0) = P(x[0] \geq \gamma; \mathcal{H}_0) = \int_{\gamma}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt = Q(\gamma)$
2. Probability of detection:  $P_D = P(\mathcal{H}_1; \mathcal{H}_1) = 1 - P_M = 1 - P(\mathcal{H}_0; \mathcal{H}_1)$
3.  $P_D = P(\mathcal{H}_1; \mathcal{H}_1) = \int_{\gamma}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(t-1)^2}{2}\right) dt = Q(\gamma - 1)$



S. Kay – detection theory Figs. 3.2 and 3.3.



0/0

Question st

### Q3: Where are the false alarm area and detection area?

II, III+IV

0%

III, III+IV

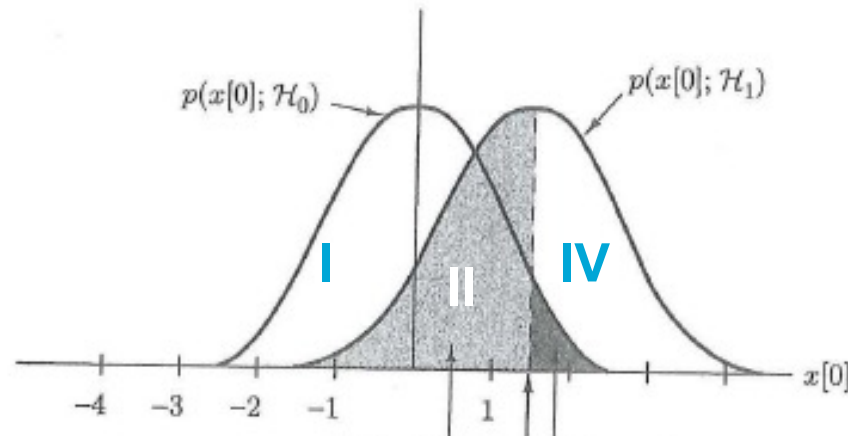
0%

II, IV

0%

III, IV

0%





0/0

### Q3: Where are the false alarm area and detection area?

II, III+IV

0%

III, III+IV



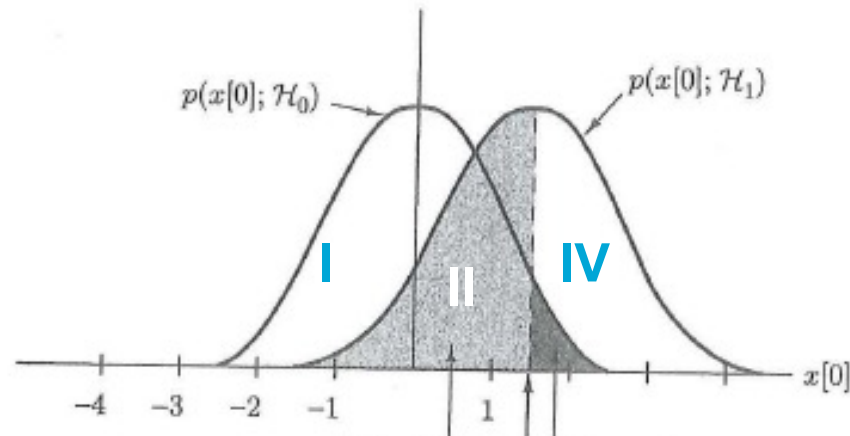
0%

II, IV

0%

III, IV

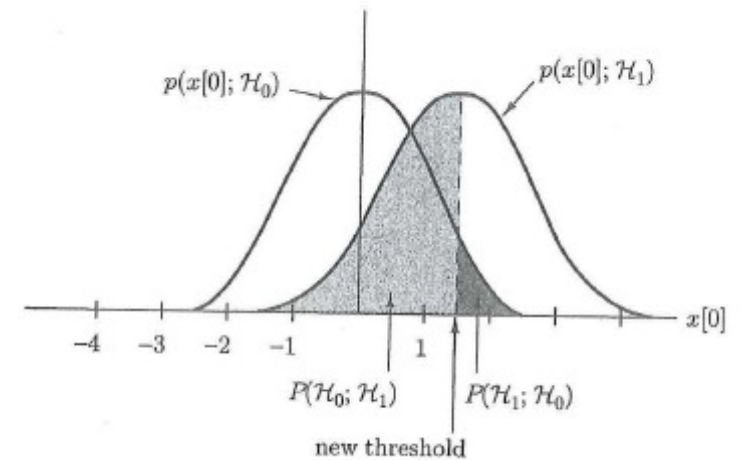
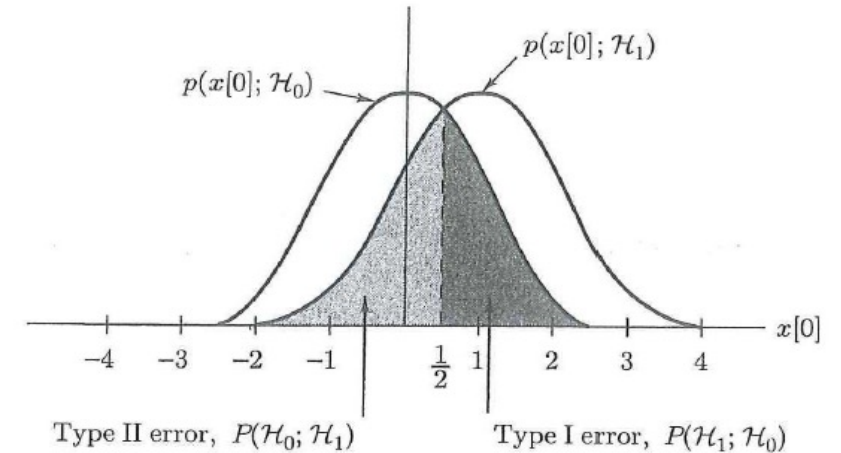
0%



Join at: vevox.

# Neyman-Pearson Theorem – Detection Performance

- These two errors can be traded off against each other.
- It is not possible to reduce both error probabilities.
- False alarm probability  $P_{FA} = P(\mathcal{H}_1; \mathcal{H}_0)$
- Detection probability  $P_D = P(\mathcal{H}_1; \mathcal{H}_1)$
- To design the optimal detector, the Neyman-Pearson approach is to maximise  $P_D$  while keeping  $P_{FA}$  fixed (small).



S. Kay – detection theory Figs. 3.2 and 3.3.

# Neyman-Pearson Theorem

- Given data set data set  $\mathbf{x} = [x[0], x[1], \dots, x[N - 1]]^T$
- Detection problem

$$\mathcal{H}_0 : T(\mathbf{x}) < \gamma$$

$$\mathcal{H}_1 : T(\mathbf{x}) > \gamma$$

1. Design a decision function  $T(\mathbf{x})$  to maximize  $P_D$  subject to  $P_{FA} < \alpha$ .
2. Determine threshold  $\gamma$ .

# Neyman-Pearson Theorem

To maximize  $P_D$  for a given  $P_{FA} = \alpha$  decide  $\mathcal{H}_1$  if

$$L(\mathbf{x}) = \frac{p(\mathbf{x}; \mathcal{H}_1)}{p(\mathbf{x}; \mathcal{H}_0)} > \lambda$$

where the threshold  $\lambda$  is found from

$$P_{FA} = \int_{\{\mathbf{x}: L(\mathbf{x}) > \lambda\}} p(\mathbf{x}; \mathcal{H}_0) d\mathbf{x} = \alpha$$

The function  $L(\mathbf{x})$  is called the likelihood ratio and the entire test is called the likelihood ratio test (LRT).

Full derivation in Appendix 3A.

# Neyman-Pearson Theorem - Derivation

$$\max P_D \text{ subject to } P_{FA} = \alpha$$

Constraint optimization, use Lagrangian:

$$\begin{aligned} F &= P_D + \lambda(P_{FA} - \alpha) \\ &= \int_{R_1} p(\mathbf{x}; \mathbf{H}_1) d\mathbf{x} + \lambda \left( \int_{R_1} p(\mathbf{x}; \mathbf{H}_0) d\mathbf{x} - \alpha \right) \\ &= \int_{R_1} (p(\mathbf{x}; \mathbf{H}_1) + \lambda p(\mathbf{x}; \mathbf{H}_0)) d\mathbf{x} - \lambda\alpha \end{aligned}$$

The problem now is (see Figures) to select to right range  $R_1$  and  $R_0$ . As we want to maximise  $F$ , a value  $\mathbf{x}$  should only be included in  $R_1$  if it increases the integrand. So,  $\mathbf{x}$  should only be included in  $R_1$  if

$$p(\mathbf{x}; \mathbf{H}_1) + \lambda p(\mathbf{x}; \mathbf{H}_0) > 0$$

# Neyman-Pearson Theorem - Derivation

$$p(\mathbf{x}; \mathbf{H}_1) + \lambda p(\mathbf{x}; \mathbf{H}_0) > 0$$

$$\Rightarrow \frac{p(\mathbf{x}; \mathbf{H}_1)}{p(\mathbf{x}; \mathbf{H}_0)} > -\lambda$$

A likelihood ratio is always positive, so  $-\lambda = \gamma > 0$  (if  $\lambda > 0$  we would have  $P_{FA} = 1$ )

$$\frac{p(\mathbf{x}; \mathbf{H}_1)}{p(\mathbf{x}; \mathbf{H}_0)} > \gamma,$$

where  $\gamma$  is found from  $P_{FA} = \alpha$ .



# Neyman-Pearson Theorem

To maximize  $P_D$  for a given  $P_{FA} = \alpha$  decide  $\mathcal{H}_1$  if

$$L(\mathbf{x}) = \frac{p(\mathbf{x}; \mathcal{H}_1)}{p(\mathbf{x}; \mathcal{H}_0)} > \lambda$$

where the threshold  $\lambda$  is found from

$$P_{FA} = \int_{\{\mathbf{x}: L(\mathbf{x}) > \lambda\}} p(\mathbf{x}; \mathcal{H}_0) d\mathbf{x} = \alpha$$

The function  $L(\mathbf{x})$  is called the likelihood ratio and the entire test is called the likelihood ratio test (LRT).

Full derivation in Appendix 3A.

Important! We thus assume that  $p(\mathbf{x}; \mathcal{H}_1)$  and  $p(\mathbf{x}; \mathcal{H}_0)$  are known!

- When the shape of  $p(\mathbf{x}; \mathcal{H}_1)$  and  $p(\mathbf{x}; \mathcal{H}_0)$  is known, but their distributional parameters are unknown: Generalized likelihood ratio test, lec. 13.
- If the  $p(\mathbf{x}; \mathcal{H}_1)$  and  $p(\mathbf{x}; \mathcal{H}_0)$  are unknown  $\rightarrow$  (Machine) Learning from data.



0/0

Question sl

## Q4: What is the role of the likelihood ratio in Neyman-Pearson hypothesis testing?

Determines the p-value

0%

Forms the basis for decision making

0%

Indicates the sample size

0%

Measures the effect size

0%

Join at: [vevox.app](https://vevox.app) ID: 135-988-786



0

## Q4: What is the role of the likelihood ratio in Neyman-Pearson hypothesis testing?

Determines the p-value

0%

Forms the basis for decision making



0%

Indicates the sample size

0%

Measures the effect size

0%

# Neyman-Pearson Theorem – Example DC in WGN

Consider the following signal detection problem

$$\mathcal{H}_0 : x[n] = w[n] \quad n = 0, 1, \dots, N - 1$$

$$\mathcal{H}_1 : x[n] = s[n] + w[n] \quad n = 0, 1, \dots, N - 1$$

where the signal is  $s[n] = A$  for  $A > 0$  and  $w[n]$  is WGN with variance  $\sigma^2$ . Now the NP detector decides  $\mathcal{H}_1$  if

# Neyman-Pearson Theorem – Example DC in WGN

Consider the following signal detection problem

$$\mathcal{H}_0 : x[n] = w[n] \quad n = 0, 1, \dots, N - 1$$

$$\mathcal{H}_1 : x[n] = s[n] + w[n] \quad n = 0, 1, \dots, N - 1$$

where the signal is  $s[n] = A$  for  $A > 0$  and  $w[n]$  is WGN with variance  $\sigma^2$ . Now the NP detector decides  $\mathcal{H}_1$  if

$$L(\mathbf{x}) = \frac{\frac{1}{(2\pi\sigma^2)^{\frac{N}{2}}} \exp \left[ -\frac{1}{2\sigma^2} \sum_{n=0}^{N-1} (x[n] - A)^2 \right]}{\frac{1}{(2\pi\sigma^2)^{\frac{N}{2}}} \exp \left[ -\frac{1}{2\sigma^2} \sum_{n=0}^{N-1} x^2[n] \right]} > \lambda$$

Taking the logarithm of both sides and simplification results in

$$T(\mathbf{x}) = \frac{1}{N} \sum_{n=0}^{N-1} x[n] > \frac{\sigma^2}{NA} \ln \lambda + \frac{A}{2} = \lambda'$$

A

$$T(\mathbf{x}) \sim \begin{cases} \mathcal{N}(0, \sigma^2) & \text{under } \mathcal{H}_0 \\ \mathcal{N}(A, \sigma^2) & \text{under } \mathcal{H}_1 \end{cases}$$

B

$$T(\mathbf{x}) \sim \begin{cases} \mathcal{N}(0, \frac{\sigma^2}{N}) & \text{under } \mathcal{H}_0 \\ \mathcal{N}(A, \frac{\sigma^2}{N}) & \text{under } \mathcal{H}_1 \end{cases}$$

C

$$T(\mathbf{x}) \sim \begin{cases} \mathcal{N}(A, \frac{\sigma^2}{2}) & \text{under } \mathcal{H}_0 \\ \mathcal{N}(0, \frac{\sigma^2}{N}) & \text{under } \mathcal{H}_1 \end{cases}$$

# Neyman-Pearson Theorem – Example DC in WGN

The NP detector compares the sample mean  $\bar{x} = \frac{1}{N} \sum_{n=0}^{N-1} x[n]$  to a threshold  $\lambda'$ . To determine the detection performance, we first note that the test statistic  $T(\mathbf{x}) = \bar{x}$  is Gaussian under each hypothesis and its distribution is as follows

$$T(\mathbf{x}) \sim \begin{cases} \mathcal{N}(0, \frac{\sigma^2}{N}) & \text{under } \mathcal{H}_0 \\ \mathcal{N}(A, \frac{\sigma^2}{N}) & \text{under } \mathcal{H}_1 \end{cases}$$

We have then  $P_{FA} = Pr(T(\mathbf{x}) > \lambda'; \mathcal{H}_0) = Q\left(\frac{\lambda'}{\sqrt{\sigma^2/N}}\right) \Rightarrow \lambda' = \sqrt{\frac{\sigma^2}{N}} Q^{-1}(P_{FA})$  and

$$P_D = Pr(T(\mathbf{x}) > \lambda'; \mathcal{H}_1) = Q\left(\frac{\lambda' - A}{\sqrt{\sigma^2/N}}\right)$$

$P_D$  and  $P_{FA}$  are related to each other according to the following equation

$$P_D = Q\left(Q^{-1}(P_{FA}) - \sqrt{\frac{NA^2}{\sigma^2}}\right) \rightarrow \text{Signal energy-to-noise ratio}$$

# Neyman-Pearson Theorem – Example DC in WGN

$$P_D = Q \left( Q^{-1}(P_{FA}) - \sqrt{\frac{NA^2}{\sigma^2}} \right)$$

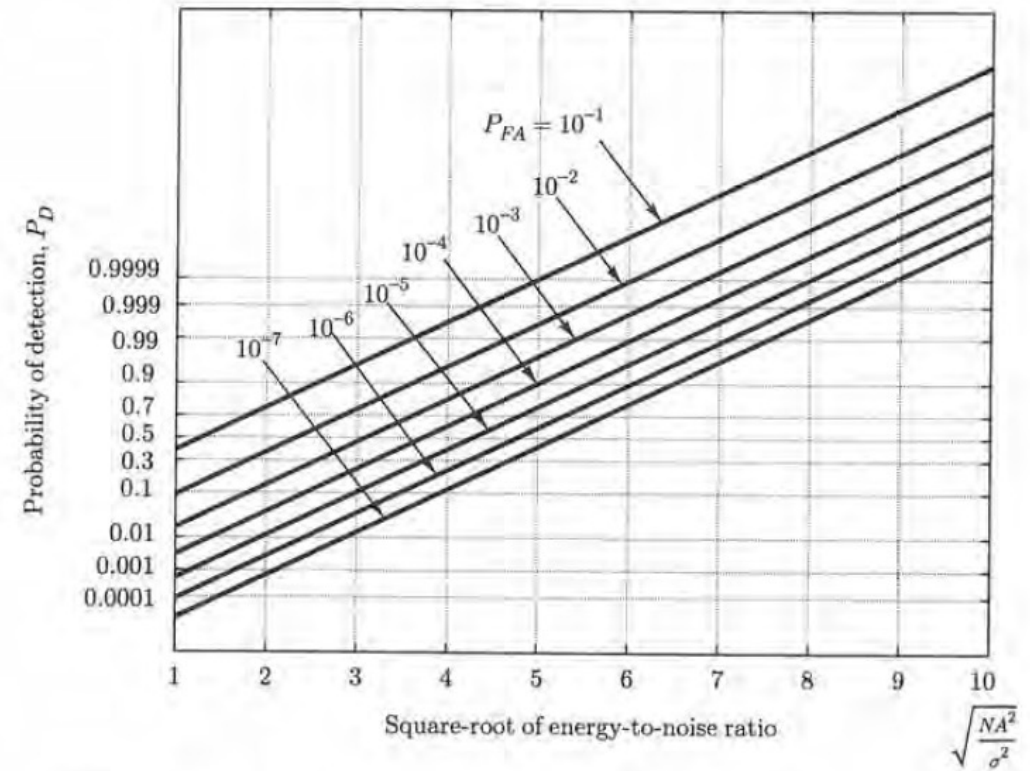
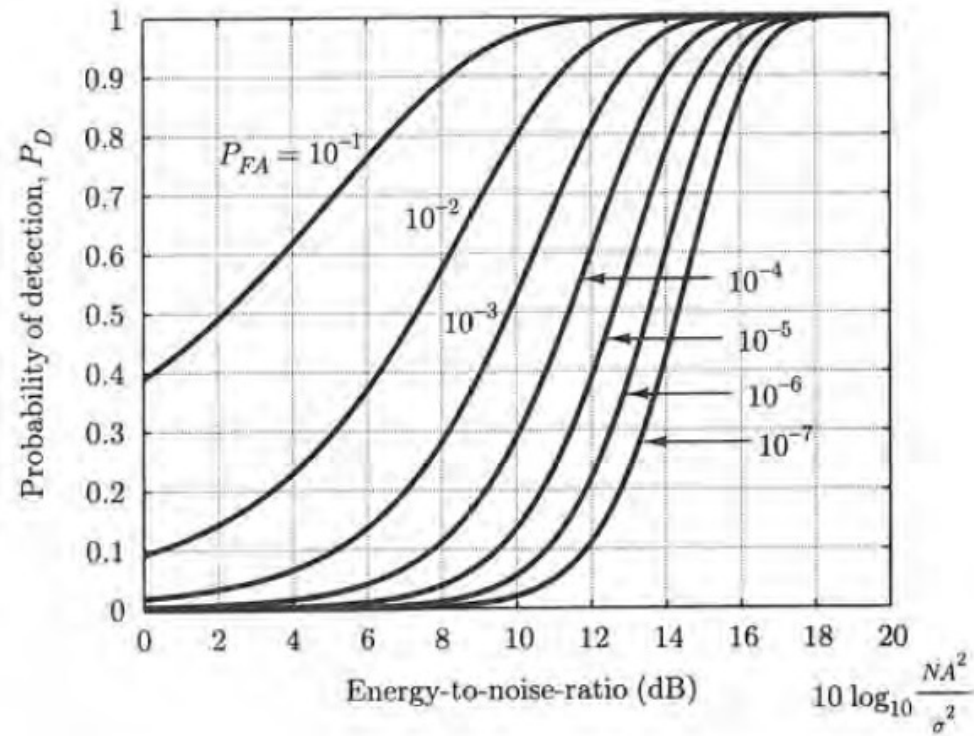
Remember the deflection coefficient

$$d^2 = \frac{(E(T; \mathcal{H}_1) - E(T; \mathcal{H}_0))^2}{\text{var}(T; \mathcal{H}_0)}$$

In this case  $d^2 = \frac{NA^2}{\sigma^2}$ .

Further notice that the detection performance ( $P_D$ ) increases monotonic with the deflection coefficient.

# Neyman-Pearson Theorem – Example DC in WGN



S. Kay – detection theory.





0/0

Question st

## Q6: Where is the False Alarm area?

A. I

0%

B. II

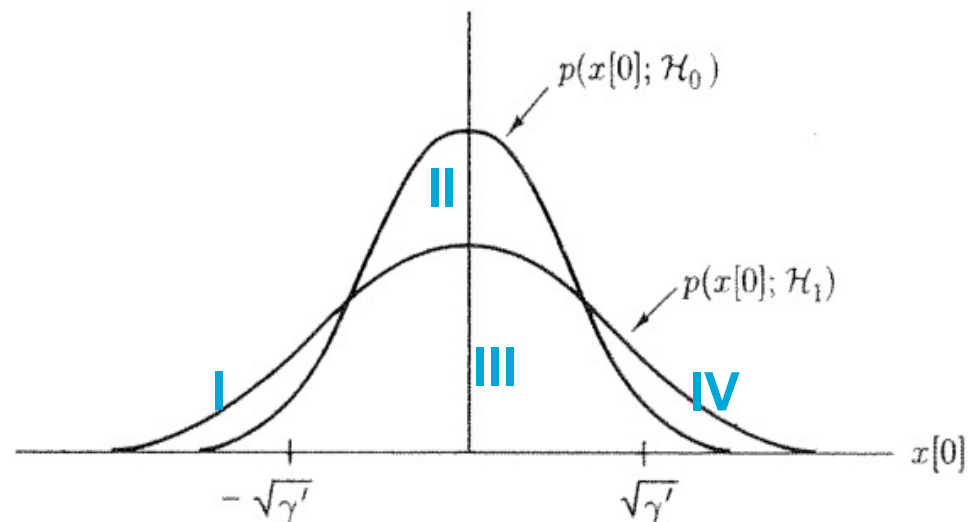
0%

C. III

0%

D. IV

0%



ID: 135-988-786

Join at: [vevox.app](https://vevox.app)



0/0

Preparing Re

## Q6: Where is the False Alarm area?

A. I

0%

B. II

0%

C. III



0%

D. IV

0%

Join at: [vevox.app](https://vevox.app) ID: 135-988-786

# NP Theorem – Example Change Var

Consider an IID process  $x[n]$ .

$$\mathcal{H}_0 : x[n] \sim \mathcal{N}(0, \sigma_0^2)$$

$$\mathcal{H}_1 : x[n] \sim \mathcal{N}(0, \sigma_1^2),$$

with  $\sigma_1^2 > \sigma_0^2$ .

Neyman-Pearson test:

$$\frac{\frac{1}{(2\pi\sigma_1^2)^{\frac{N}{2}}} \exp\left[-\frac{1}{2\sigma_1^2} \sum_{n=0}^{N-1} x[n]^2\right]}{\frac{1}{(2\pi\sigma_0^2)^{\frac{N}{2}}} \exp\left[-\frac{1}{2\sigma_0^2} \sum_{n=0}^{N-1} x^2[n]\right]} > \gamma$$

What is

- What is  $T(\mathbf{x})$ ?
- The distribution of  $T(\mathbf{x})$  under both  $\mathcal{H}_0$  and  $\mathcal{H}_1$ ?
- The optimal threshold  $\lambda'$ ?
- $P_D$ ?

A

B

C

$$T(\mathbf{x}) = \frac{1}{N} \sum_{n=0}^{N-1} |x[n]|$$

$$T(\mathbf{x}) = \frac{1}{N} \sum_{n=0}^{N-1} x[n]$$

$$T(\mathbf{x}) = \frac{1}{N} \sum_{n=0}^{N-1} x^2[n]$$

# NP Theorem – Example Change Var

A

$$T(\mathbf{x}) = \frac{1}{N} \sum_{n=0}^{N-1} |x[n]|$$

B

$$T(\mathbf{x}) = \frac{1}{N} \sum_{n=0}^{N-1} x[n]$$

C

$$T(\mathbf{x}) = \frac{1}{N} \sum_{n=0}^{N-1} x^2[n]$$

# NP Theorem – Example Change Var

Taking the logarithm of both sides and simplification results in

$$-\frac{1}{2} \left( \frac{1}{\sigma_1^2} - \frac{1}{\sigma_0^2} \right) \sum_{n=0}^{N-1} x^2[n] > \ln \gamma + \frac{N}{2} \ln \frac{\sigma_1^2}{\sigma_0^2}$$

we then have

$$T(\mathbf{x}) = \frac{1}{N} \sum_{n=0}^{N-1} x^2[n] > \gamma'$$

$$\text{with } \gamma' = \frac{\frac{2}{N} \ln \gamma + \ln \frac{\sigma_1^2}{\sigma_0^2}}{\frac{1}{\sigma_0^2} - \frac{1}{\sigma_1^2}}$$

What is

- What is  $T(\mathbf{x})$ ?
- The distribution of  $T(\mathbf{x})$  under both  $\mathcal{H}_0$  and  $\mathcal{H}_1$ ?
- The optimal threshold  $\lambda'$ ?
- $P_D$ ?

# NP Theorem – Example Change Var

For  $N = 1$  we decide for  $\mathcal{H}_1$  if:

$$|x[0]| > \sqrt{\gamma'}.$$

$$P_{FA} = Pr \left\{ |x[0]| > \sqrt{\gamma'}; \mathcal{H}_0 \right\} = 2Pr \left\{ x[0] > \sqrt{\gamma'}; \mathcal{H}_0 \right\}.$$

$$P_{FA} = 2Q \left( \frac{\sqrt{\gamma'}}{\sqrt{\sigma_0^2}} \right).$$

$$\Rightarrow \sqrt{\gamma'} = Q^{-1} \left( \frac{1}{2} P_{FA} \right) \sqrt{\sigma_0^2}$$

$$P_D = 2Q \left( \frac{Q^{-1} \left( \frac{1}{2} P_{FA} \right) \sqrt{\sigma_0^2}}{\sqrt{\sigma_1^2}} \right)$$

# NP Theorem – Example Change Var

A

$$T(\mathbf{x}) = \begin{cases} \frac{\sum_{n=0}^{N-1} x^2[n]}{\sigma_0^2} \sim \chi_N^2 & \text{under } \mathcal{H}_0 \\ \frac{\sum_{n=0}^{N-1} x^2[n]}{\sigma_1^2} \sim \chi_N^2 & \text{under } \mathcal{H}_1 \end{cases}$$

B

$$T(\mathbf{x}) = \begin{cases} \frac{\sum_{n=0}^{N-1} x^2[n]}{\sigma_1^2} \sim \chi_N^2 & \text{under } \mathcal{H}_0 \\ \frac{\sum_{n=0}^{N-1} x^2[n]}{\sigma_0^2} \sim \chi_N^2 & \text{under } \mathcal{H}_1 \end{cases}$$

C

$$T(\mathbf{x}) \sim \begin{cases} \mathcal{N}(0, \sigma_0^2) & \text{under } \mathcal{H}_0 \\ \mathcal{N}(0, \sigma_1^2) & \text{under } \mathcal{H}_1 \end{cases}$$

# NP Theorem – Example Change Var

$$T(\mathbf{x}) = \begin{cases} \frac{\sum_{n=0}^{N-1} x^2[n]}{\sigma_0^2} \sim \chi_N^2 & \text{under } \mathcal{H}_0 \\ \frac{\sum_{n=0}^{N-1} x^2[n]}{\sigma_1^2} \sim \chi_N^2 & \text{under } \mathcal{H}_1 \end{cases}$$

For general  $N$  we decide for  $\mathcal{H}_1$  if:

$$\sum_{n=0}^{N-1} x^2[n] > \gamma'.$$

$$P_{FA} = Pr \left\{ \sum_{n=0}^{N-1} x^2[n] > \gamma'; \mathcal{H}_0 \right\} = Pr \left\{ \frac{\sum_{n=0}^{N-1} x[n]^2}{\sigma_0^2} > \frac{\gamma'}{\sigma_0^2}; \mathcal{H}_0 \right\}.$$

$$P_{FA} = Q_{\chi_N^2} \left( \frac{\gamma'}{\sigma_0^2} \right).$$

$$\Rightarrow \gamma' = Q_{\chi_N^2}^{-1}(P_{FA}) \sigma_0^2$$

$$P_D = Q_{\chi_N^2} \left( \frac{Q_{\chi_N^2}^{-1}(P_{FA}) \sigma_0^2}{\sigma_1^2} \right)$$



# Minimum Probability of Error

Assume the prior probabilities of  $\mathcal{H}_0$  and  $\mathcal{H}_1$  are known and represented by  $P(\mathcal{H}_0)$  and  $P(\mathcal{H}_1)$ , respectively. The probability of error,  $P_e$ , is then defined as

$$P_e = P(\mathcal{H}_1)P(\mathcal{H}_0|\mathcal{H}_1) + P(\mathcal{H}_0)P(\mathcal{H}_1|\mathcal{H}_0) = P(\mathcal{H}_1)P_M + P(\mathcal{H}_0)P_{FA}$$

Our goal is to design a detector that minimizes  $P_e$ . It is shown that the following detector is optimal in this case

$$\frac{p(\mathbf{x}|\mathcal{H}_1)}{p(\mathbf{x}|\mathcal{H}_0)} > \frac{P(\mathcal{H}_0)}{P(\mathcal{H}_1)} = \lambda$$

Derivation in Appendix 3B.

# Minimum Probability of Error - Derivation

$$\begin{aligned} P_e &= P(\mathcal{H}_1)P(\mathcal{H}_0|\mathcal{H}_1) + P(\mathcal{H}_0)P(\mathcal{H}_1|\mathcal{H}_0) \\ &= P(\mathcal{H}_1) \int_{R_0} p(\mathbf{x}|\mathcal{H}_1) d\mathbf{x} + P(\mathcal{H}_0) \int_{R_1} p(\mathbf{x}|\mathcal{H}_0) d\mathbf{x} \end{aligned}$$

We know that

$$\int_{R_0} p(\mathbf{x}|\mathcal{H}_1) d\mathbf{x} = 1 - \int_{R_1} p(\mathbf{x}|\mathcal{H}_1) d\mathbf{x},$$

such that

$$\begin{aligned} P_e &= P(\mathcal{H}_1) \left( 1 - \int_{R_1} p(\mathbf{x}|\mathcal{H}_1) d\mathbf{x} \right) + P(\mathcal{H}_0) \int_{R_1} p(\mathbf{x}|\mathcal{H}_0) d\mathbf{x} \\ &= P(\mathcal{H}_1) + \int_{R_1} [P(\mathcal{H}_0)p(\mathbf{x}|\mathcal{H}_0) - P(\mathcal{H}_1)p(\mathbf{x}|\mathcal{H}_1)] d\mathbf{x} \end{aligned}$$

# Minimum Probability of Error - Derivation

$$P_e = P(\mathcal{H}_1) + \int_{\mathcal{R}_1} [P(\mathcal{H}_0)p(\mathbf{x}|\mathcal{H}_0) - P(\mathcal{H}_1)p(\mathbf{x}|\mathcal{H}_1)] d\mathbf{x}$$

We want to minimize  $P_e$ , so an  $\mathbf{x}$  should only be included in the region  $\mathcal{R}$  if the integrand

$$[P(\mathcal{H}_0)P(\mathbf{x}|\mathcal{H}_0) - P(\mathcal{H}_1)P(\mathbf{x}|\mathcal{H}_1)]$$

is negative for that  $\mathbf{x}$ .

$$\begin{aligned} P(\mathcal{H}_0)p(\mathbf{x}|\mathcal{H}_0) &< P(\mathcal{H}_1)p(\mathbf{x}|\mathcal{H}_1) \\ \frac{p(\mathbf{x}|\mathcal{H}_1)}{p(\mathbf{x}|\mathcal{H}_0)} &> \frac{P(\mathcal{H}_0)}{P(\mathcal{H}_1)} = \gamma \end{aligned}$$

# Minimum Probability of Error– Example DC in WGN

Consider the following signal detection problem

$$\mathcal{H}_0 : x[n] = w[n] \quad n = 0, 1, \dots, N - 1$$

$$\mathcal{H}_1 : x[n] = s[n] + w[n] \quad n = 0, 1, \dots, N - 1$$

where the signal is  $s[n] = A$  for  $A > 0$  and  $w[n]$  is WGN with variance  $\sigma^2$ . Now the min. probability of error detector decides  $\mathcal{H}_1$  if  $\frac{p(\mathbf{x}|\mathcal{H}_1)}{p(\mathbf{x}|\mathcal{H}_0)} > \frac{P(\mathcal{H}_0)}{P(\mathcal{H}_1)} = 1$  (assuming  $P(\mathcal{H}_0) = P(\mathcal{H}_1) = 0.5$ ), leading to

$$\frac{\frac{1}{(2\pi\sigma^2)^{\frac{N}{2}}} \exp\left[-\frac{1}{2\sigma^2} \sum_{n=0}^{N-1} (x[n] - A)^2\right]}{\frac{1}{(2\pi\sigma^2)^{\frac{N}{2}}} \exp\left[-\frac{1}{2\sigma^2} \sum_{n=0}^{N-1} x^2[n]\right]} > 1$$

Taking the logarithm of both sides and simplification results in

$$\frac{1}{N} \sum_{n=0}^{N-1} x[n] > \frac{A}{2}$$

# Minimum Probability of Error– Example DC in WGN

$P_e$  is then given by

$$\begin{aligned} P_e &= \frac{1}{2} [P(\mathcal{H}_0|\mathcal{H}_1) + P(\mathcal{H}_1|\mathcal{H}_0)] \\ &= \frac{1}{2} \left[ Pr\left(\frac{1}{N} \sum_{n=0}^{N-1} x[n] < A/2 | \mathcal{H}_1\right) + Pr\left(\frac{1}{N} \sum_{n=0}^{N-1} x[n] > A/2 | \mathcal{H}_0\right) \right] \\ &= \frac{1}{2} \left[ \left(1 - Q\left(\frac{A/2 - A}{\sqrt{\sigma^2/N}}\right)\right) + Q\left(\frac{A/2}{\sqrt{\sigma^2/N}}\right) \right] \\ &= Q\left(\sqrt{\frac{NA^2}{4\sigma^2}}\right) \end{aligned}$$

# Reading Tasks

- Chapter 2: PDFs:
  - Gaussian, Chi-squared, F, Rayleigh
- Study Chapter 3-3.4, 3.6-3.8
- Appendix 3A: Follow derivation of the Neyman-Pearson lemma
- Appendix 3B: follow derivation of the Minimum Probability of Error ( $P_e$ )
  
- Try solving problem 1 (see slides at the end)

# Next Lecture

## Lecture 2: Deterministic Signals (Ch. 4)

- Detection of deterministic signals:
  - Detecting a known signal in noise using the NP criterion.
- Detection of realisations of random processes :
  - Detecting the realisation of randomprocess with known covariance structure.