STATISTICAL SIGNAL PROCESSING

$\boxed{\text{Statistical DSP}}$

Digital $\longrightarrow$ discrete-time, sampled, quantized

Signal $\longrightarrow$ waveform, sequence of measurements or observations

Processing $\longrightarrow$ analyze, modify, synthesize

Examples of digital signals

- sampled speech waveform
- pixelized image
- Dow-Jones index
- stream of Internet packets
- vector of medical predictors

What kind of processing might be desirable?

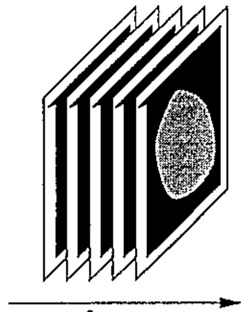
A major difficulty

In many DSP applications, we don't have complete or perfect knowledge of the signals we wish to process. We are faced with many <u>unknowns</u> and <u>uncertainties.</u>

Examples

- unknown signal parameters
  - delay of radar return
  - pitch of speech signal
- environmental noise
  - multipath signals in wireless comm.
  - ambient EM waves
  - radar jamming
- sensor noise
  - grainy images
  - old phonograph recordings
- variability inherent in nature
  - stock market
  - Internet

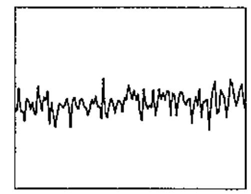
Functional Magnetic Resonance Imaging



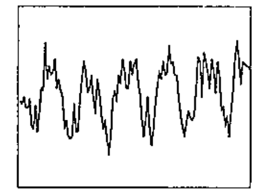
time

brain image
time-series

$\Rightarrow$



"inactive" pixel
time-series

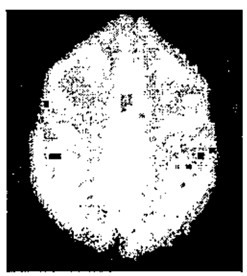


"active" pixel
time-series

$\Downarrow$

$\boxed{\text{statistical signal processing}}$

$\Leftarrow$



activation map

<u>Challenges:</u>

- measurement noise
- intrinsic uncertainties in signal behavior

How can we process signals in the face of such uncertainty?

Can we model the uncertainty and incorporate this model into the processing?

Statistical signal processing is the study of these questions.

Modeling uncertainty

The most widely accepted and commonly used approach to modeling uncertainty

(a) is ________________, although alternatives exist such as fuzzy logic.

Probability theory models uncertainty by specifying the chance of observing certain signals.

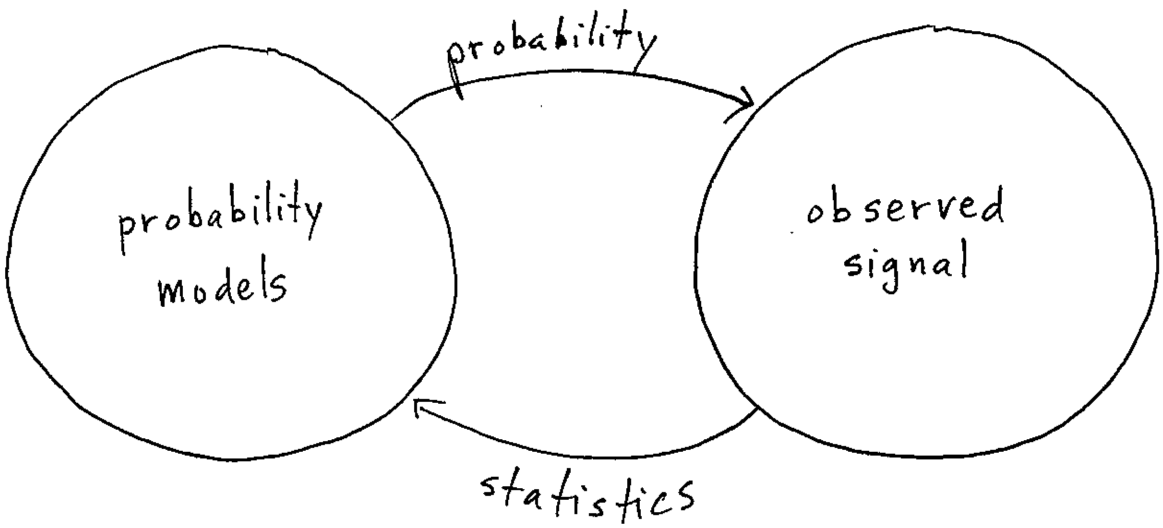
Alternatively, one can view probability as specifying the degree to which we <u>believe</u> a signal reflects the true <u>state of nature</u>.

<u>Examples</u> of probabilistic models

- sensor noise modeled as an additive <u>Gaussian</u> random variable
- uncertainty in the phase of a sinusoidal signal modeled as a ______ random variable on $[0, 2\pi)$. (b)
- uncertainty in the number of photons striking a CCD per unit time modeled as a ______ random variable. (c)

Probability laws describe the uncertainty in the signals we <u>might</u> observe.

Statistics describe the salient features of the signals we <u>do</u> observe, and allow us to draw conclusions (inferences) about which probability model actually reflects the true state of nature



Statistical inference

A <u>statistic</u> is a function of observed data, and may be scalar or vector valued.

<u>Examples</u> Suppose we observe n scalar values $x_1, \dots, x_N$. The following are statistics

- sample mean $\longrightarrow$ $\bar{x} = \frac{1}{N} \sum_{i=1}^{N} x_i$
- the data itself $\longrightarrow$ $[x_1, \cdots, x_N]^T$
- an order statistic $\longrightarrow$ $x_{(1)} = \min\{x_1, \dots, x_N\}$
- an arbitrary function $\longrightarrow$ $[x_1^2 - x_2 \sin(x_3), e^{-x_1 x_3}]^T$

A statistic <u>cannot</u> depend on unknown quantities.

Statistical signal processing (in a nutshell)

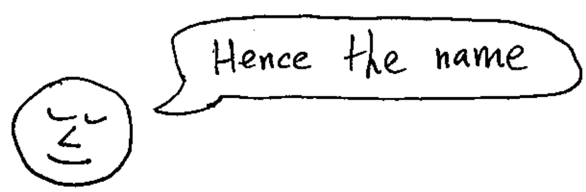
Step 1: Postulate a probability model (or models) that can be expected to reasonably capture the uncertainties in the data.

Step 2: Collect data

Step 3: Formulate statistics that allow us to interpret or understand our probability models.

In this class we will focus on three areas:

- Estimation
- Filtering
- Detection



Estimation

If our probability model has free parameters, can we use the measured signal to infer the actual parameter values?

Examples

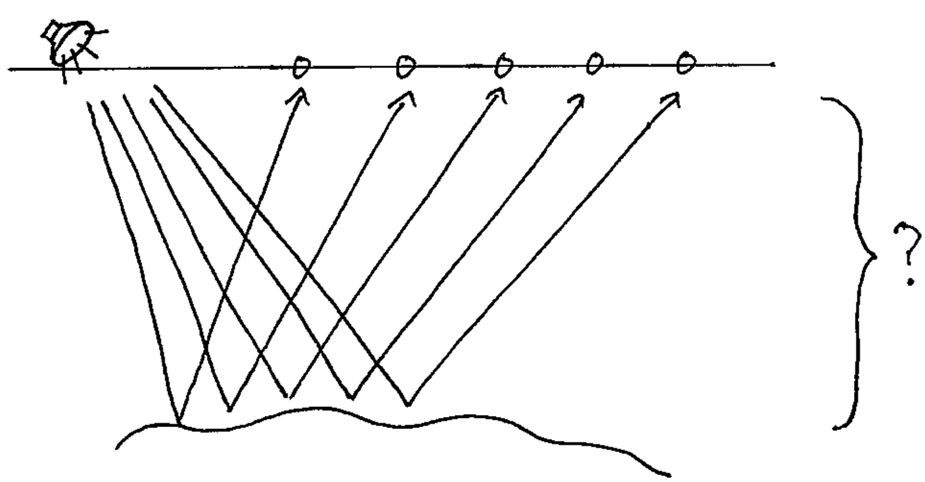
- Signal denoising: If we observe

$$\underline{x} = \underline{s} + \underline{w}$$

where $\underline{s}$ is a signal of interest and $\underline{w}$ is noise, estimate $\underline{s}$.

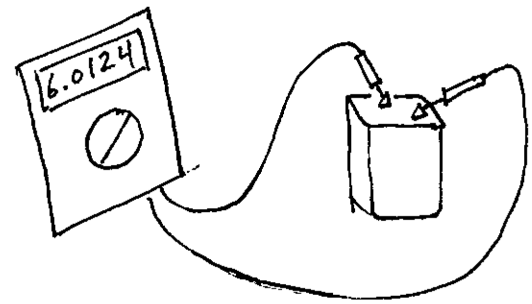
- If $s(n) = A\cos(2\pi f n + \phi)$, estimate A, ϕ, f in signal plus noise model.

- Seismology: estimate depth below ground of an oil pool based on reflected acoustic waves.



Here is a more concrete example:

<u>Example</u> Suppose we measure the voltage A of a battery using a voltmeter. Because the voltmeter tends to pick up noise from nearby objects, we take n measurements in hopes of gaining some accuracy.



Step 1: Assume a Gaussian noise model

$$x_i = A + w_i, \quad i = 1, \ldots, N$$

where

$$w_i \sim N(0, \sigma_w^2)$$

Step 2: Gather data

Step 3: Estimate A via the sample mean

$$\hat{A} = \bar{x} = \frac{1}{N} \sum_{i=1}^{N} x_i$$

Is this the "right" statistic for this noise model? How accurate is the estimate? What if σ_w^2 is unknown?

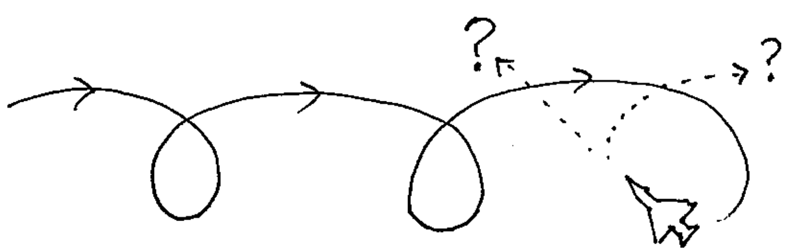
Filtering

Filtering in deterministic DSP refers to modifying a signal by means of a linear operator (usually expressed in terms of convolution).

In statistical DSP, filtering refers to signal estimation by means of a linear function of the data. Thus filtering is a special case of estimation, but it is distinguished by being concerned with online/real-time estimation of <u>streaming data</u>.

Examples

- Denoise a speech/audio signal real-time
- Track a moving target or predict its future location



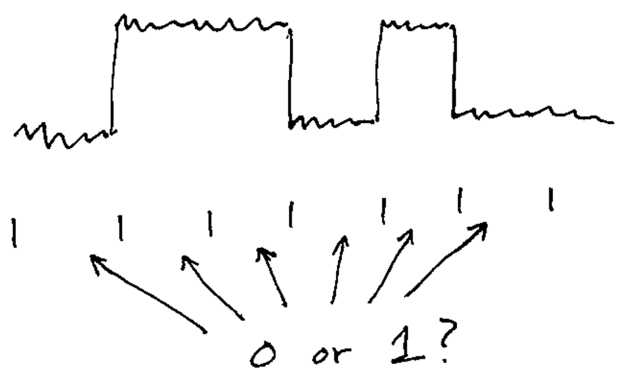
Detection

Given two (or more) probability models, which one best explains the observed signal?

Alternatively, given a single probability model, is it or is it not a valid characterization of the data?

Examples

- Decode a comm. signal into a sequence of 0's and 1's.



- Are other ships present on radar/sonar, and if so, are they friendly?
- Is a supernova exploding somewhere in the field of view of a certain telescope?

A more concrete detection example.

Example 1 Suppose you are given a coin and are asked to determine whether the following hypothesis is true:

$$H_0 : \text{the coin is fair}$$

Step 1: Assume each toss of the coin is a realization of a ______ random var.

(d)

$$X \sim$$

where $p = \text{Prob}\{\text{heads}\}$.

Step 2: Toss the coin $n = 100$ times

$$x_i = 1 \iff \text{heads}$$

$$x_i = 0 \iff \text{tails}$$

Step 3: To assess the hypothesis

$$H_0 : p = \tfrac{1}{2},$$

form the statistic

$$k = \sum_{i=1}^{100} x_i$$

and reject H_0 if $|k - 50| > 10$.

In these examples, we used our intuition and heuristics in Step 3 to solve estimation and detection problems.

In this course we will develop principled and mathematically rigorous approaches to estimation, filtering, and detection using the theoretical framework of probability and statistics.

Summary

DSP = processing digital signals with computer algorithms

SSP = statistical DSP

= processing in the presence of uncertainties and unknowns.

Key

a. probability theory

b. uniform

c. Poisson

d. Bernoulli