

# INF 564- Probability & Statistics for Data Science

Why do (electrical and Computer)

Engineers study probability?

- Natural Phenomena are uncertain and hard to be modeled precisely
- As a result, the data that are

Collected are uncertain

- Probability theory is the most important and useful theory that analyzes uncertainty
- Probability Theory has enabled

humans to build elegant

information technology solutions

- Information Technology in turn made collecting big amounts of data possible

Probability Theory, through a powerful tool called Statistics, provides a plethora of methods to analyze and model data

In some sense:

Engineering = Analysis

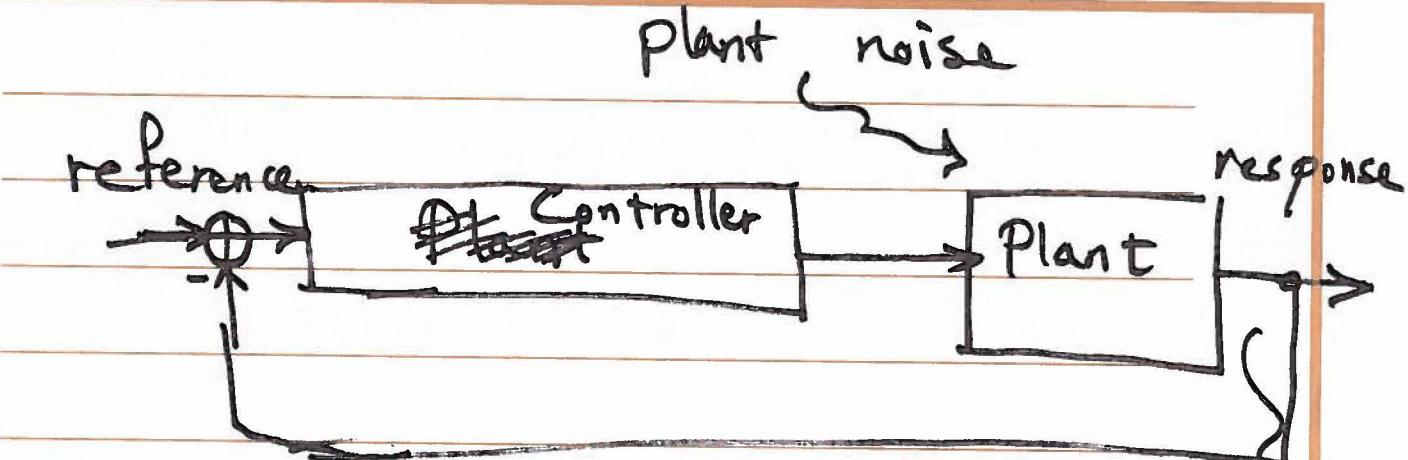
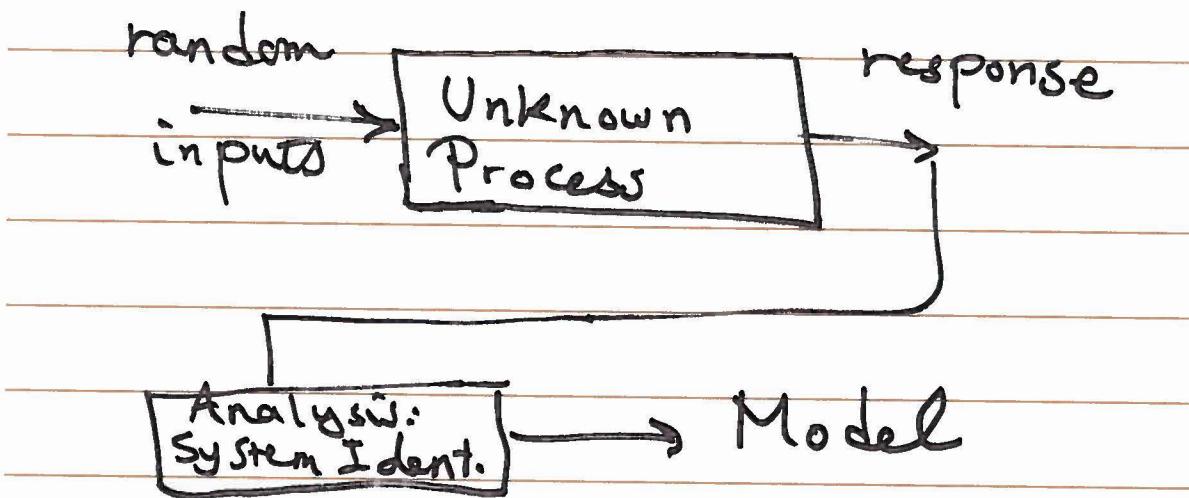
and Design Using Data

Studying Probability is a

MUST.

Examples:

## Control Systems



noise can adversely affect  
the performance of the system

Probability Theory provides

analysis tools to model

the effect of noise and to  
mitigate it

## Optical Communication



{

Interface between optical and  
electronic subsystems

Number of photoelectrons produced

modeled by a Poisson Random

Variable  $\Rightarrow$  Used to determine

if the transmitted bit was 0 or 1

# Machine Learning:

- Uses Statistics to build algorithms that learn patterns from data
- Probability and Statistics

are also used to measure how well Learning systems work

## Wireless Communications:

Analysis of the effect of

(Gaussian) noise on signals

that carry information is

at the heart of Communications

One of the most important

branches of Probability Theory,

i.e. Information Theory, gave

rise to the modern Communications

Theory. (Contributed by

Claude Shannon)

# Variability in the design of digital and analogue electronic circuits

The - frequency of  
a certain microprocessor or  
the input impedance of

an electronic amplifier may  
vary because of design  
imperfections.

Parameter estimation and Confidence  
Intervals are used to

model such variabilities by

observing an "enough" number of instances

### Computer Networks : Markov

Chains are used to model

network traffic

### Financial Engineering

Probability is vastly used to

determine the risk and expected

return of portfolios, analyze

stock market prices (through time series analysis), etc.

## Reliability and Survival

Analysis: Probability theory is heavily used to model

the Mean-Time Between Failure (MTBF) of equipments, to measure their reliability, and to ~~est~~ model their useful life.

Similar techniques are

used in determining the  
effect of a new drug on  
survival rates of (cancer)  
patients