Introduction to Random Processes

Gonzalo Mateos
Dept. of ECE and Goergen Institute for Data Science
University of Rochester
gmateosb@ece.rochester.edu
http://www.hajim.rochester.edu/ece/sites/gmateos/

August 28, 2023
Introductions

Class description and contents

Gambling
Who we are, where to find me, lecture times

- Gonzalo Mateos
- Associate Professor, Dept. of Electrical and Computer Engineering
- CSB 726, gmateosb@ece.rochester.edu
- http://www.hajim.rochester.edu/ece/sites/gmateos/
- Where? We meet in Gavett Hall 202
- When? Mondays and Wednesdays 4:50 pm to 6:05 pm
  - Due to travel, few make-up lectures on Fridays 4:50 pm to 6:05 pm
- My office hours, Tuesdays at 10:30 am
  - Anytime, as long as you have something interesting to tell me
- Class website
  http://www.hajim.rochester.edu/ece/sites/gmateos/ECE440.html
A great TA to help you with your homework

- Zhuo Liu
- CSB 527, zliu106@ur.rochester.edu
- His office hours, Fridays at 2:30 pm
Prerequisites

(I) Probability theory
- Random (Stochastic) processes are collections of random variables
- Basic knowledge expected. *Will review in the first six lectures*

(II) Calculus and linear algebra
- Integrals, limits, infinite series, differential equations
- Vector/matrix notation, systems of linear equations, eigenvalues

(III) Programming in Matlab
- Needed for homework
  https://tech.rochester.edu/software/matlab/
- If you know programming you can learn Matlab in one afternoon
  ⇒ But it has to be *one of this week’s afternoons*
Homework, exams and grading

(I) **Homework sets** (10 in 15 weeks) worth 28 points
    - Important and demanding part of this class
    - Collaboration accepted, welcomed, and encouraged

(II) **Midterm examination** on Monday **October 30** worth 36 points

(III) **Final take-home examination** on **December 17-19** worth 36 points
    - Work independently.  **This time no collaboration, no discussion**
    - ECE 271 students get 10 free points
    - **At least 60 points are required for passing** (C grade)
    - B requires at least 75 points.  **A at least 92. No curve**
      ⇒ Goal is for everyone to earn an A
Textbooks

- Good general reference for the class

  ⇒ Available online: http://www.library.rochester.edu/

- Also nice for topics including Markov chains, queuing models


- Both on reserve for the class in Carlson Library
Be nice

- I work hard for this course, expect you to do the same
- Please come to class, be on time, pay attention, ask
- Do all of your homework
- Do not hand in as yours the solution of others (or mine)
- Do not collaborate in the exams
- A little bit of (conditional) probability ...
- Probability of getting an E in this class is 0.04
- Probability of getting an E given you skip 4 homework sets is 0.7
  \[ \Rightarrow \] I’ll give you three notices, afterwards, I’ll give up on you
- Come and learn. Useful down the road
Class contents

Introductions

Class description and contents

Gambling
Stochastic systems

- **Stochastic system**: Anything random that evolves in time
  - Time can be **discrete** \( n = 0, 1, 2, \ldots \), or **continuous** \( t \in [0, \infty) \)
- More formally, random processes assign a function to a random event
- Compare with “random variable assigns a value to a random event”
- Can interpret a random process as a collection of random variables
  - Generalizes concept of **random vector to functions**
  - Or generalizes the concept of **function to random settings**
A voice recognition system

- Random event $\sim$ word spoken. Random process $\sim$ the waveform
- Try the file speech_signals.m

“Hi”

“Good”

“Bye”

‘S’
Four thematic blocks

(I) Probability theory review (6 lectures)
   ▶ Probability spaces, random variables, independence, expectation
   ▶ Conditional probability: time $n + 1$ given time $n$, future given past ...
   ▶ Limits in probability, almost sure limits: behavior as $n \to \infty$ ...
   ▶ Common probability distributions (binomial, exponential, Poisson, Gaussian)
	n▶ Random processes are complicated entities
   ⇒ Restrict attention to particular classes that are somewhat tractable

(II) Markov chains (6 lectures)

(III) Continuous-time Markov chains (7 lectures)

(IV) Stationary random processes (8 lectures)

▶ Midterm covers up to Markov chains
Probability theory is a formalism to work with uncertainty
- Given a data-generating process, what are properties of outcomes?

Statistical inference deals with the inverse problem
- Given outcomes, what can we say on the data-generating process?
- ECE409 - Machine Learning, ECE442 - Network Science Analytics, CSC440 - Data Mining, ECE441 - Detection and Estimation Theory, ...
Markov chains

- **Countable set of states** $1, 2, \ldots$. At discrete time $n$, state is $X_n$.
- **Memoryless (Markov) property**
  - Probability of next state $X_{n+1}$ depends on current state $X_n$.
  - But not on past states $X_{n-1}, X_{n-2}, \ldots$.

- Can be happy ($X_n = 0$) or sad ($X_n = 1$).
- Tomorrow’s mood only affected by today’s mood.
- Whether happy or sad today, likely to be happy tomorrow.
- But when sad, a little less likely so.
- **Of interest**: classification of states, ergodicity, limiting distributions.
- **Applications**: Google’s PageRank, communication networks, queues, reinforcement learning, ...
Continuous-time Markov chains

- **Countable set of states** $1, 2, \ldots$. **Continuous-time index** $t$, state $X(t)$
  - **Transition** between states can happen at any time
  - **Markov**: Future independent of the past given the present

- Probability of changing state in an infinitesimal time $dt$

- **Of interest**: Poisson processes, exponential distributions, transition probabilities, Kolmogorov equations, limit distributions

- **Applications**: Chemical reactions, queues, epidemic modeling, traffic engineering, weather forecasting, ...
Stationary random processes

- **Continuous** time \( t \), **continuous state** \( X(t) \), not necessarily Markov
- Prob. distribution of \( X(t) \) constant or becomes constant as \( t \) grows
  \( \Rightarrow \) System has a **steady state in a random sense**
- **Of interest**: Brownian motion, white noise, Gaussian processes, autocorrelation, power spectral density
- **Applications**: Black Scholes model for option pricing, radar, face recognition, noise in electric circuits, filtering and equalization, ...
Gambling

Introductions

Class description and contents

Gambling
An interesting betting game

- There is a certain game in a certain casino in which ...
  - Your chances of winning are $p > 1/2$

- You place $1$ bets
  - (a) With probability $p$ you gain $1$; and
  - (b) With probability $1 - p$ you lose your $1$ bet

- The catch is that you either
  - (a) Play until you go broke (lose all your money)
  - (b) Keep playing forever

- You start with an initial wealth of $w_0$

- Q: Shall you play this game?
Let $t$ be a time index (number of bets placed)

Denote as $X(t)$ the outcome of the bet at time $t$

$X(t) = 1$ if bet is won (w.p. $p$)
$X(t) = 0$ if bet is lost (w.p. $1 - p$)

$X(t)$ is called a Bernoulli random variable with parameter $p$

Denote as $W(t)$ the player’s wealth at time $t$. Initialize $W(0) = w_0$

At times $t > 0$ wealth $W(t)$ depends on past wins and losses

When bet is won $W(t + 1) = W(t) + 1$
When bet is lost $W(t + 1) = W(t) - 1$

More compactly can write $W(t + 1) = W(t) + (2X(t) - 1)$
Only holds so long as $W(t) > 0$
t = 0; w(t) = w_0; max_t = 10^3; // Initialize variables
% repeat while not broke up to time max_t
while (w(t) > 0) & (t < max_t) do
    \( x(t) = \text{random('bino',1,p)}; \) % Draw Bernoulli random variable
    if \( x(t) == 1 \) then
        \( w(t + 1) = w(t) + b; \) % If \( x = 1 \) wealth increases by \( b \)
    else
        \( w(t + 1) = w(t) - b; \) % If \( x = 0 \) wealth decreases by \( b \)
    end
    \( t = t + 1; \)
end

▶ Initial wealth \( w_0 = 20 \), bet \( b = 1 \), win probability \( p = 0.55 \)
▶ Q: Shall we play?
One lucky player

She didn’t go broke. After $t = 1000$ bets, her wealth is $W(t) = 109$

⇒ Less likely to go broke now because wealth increased
Two lucky players

- After $t = 1000$ bets, wealths are $W_1(t) = 109$ and $W_2(t) = 139$

$\Rightarrow$ Increasing wealth seems to be a pattern
Ten lucky players

- Wealths $W_j(t)$ after $t = 1000$ bets between 78 and 139

$\Rightarrow$ Increasing wealth is definitely a pattern
One unlucky player

- But this does not mean that all players will turn out as winners
  ⇒ The twelfth player $j = 12$ goes broke
One unlucky player

- But this does not mean that all players will turn out as winners
  - The twelfth player $j = 12$ goes broke
One hundred players

- All players (except for $j = 12$) end up with substantially more money
It is not difficult to find a line estimating the average of $W(t)$

$$\bar{w}(t) \approx w_0 + (2p - 1)t \approx w_0 + 0.1t \quad \text{(recall } p = 0.55\text{)}$$
Where does the average tendency come from?

- Assuming we do not go broke, we can write
  \[ W(t+1) = W(t) + \left(2X(t) - 1\right), \quad t = 0, 1, 2, \ldots \]

- The assumption is incorrect as we saw, but suffices for simplicity

- Taking expectations on both sides and using linearity of expectation
  \[ \mathbb{E}[W(t+1)] = \mathbb{E}[W(t)] + \left(2\mathbb{E}[X(t)] - 1\right) \]

- The expected value of Bernoulli \( X(t) \) is
  \[ \mathbb{E}[X(t)] = 1 \times P(X(t) = 1) + 0 \times P(X(t) = 0) = p \]

- Which yields \( \Rightarrow \mathbb{E}[W(t+1)] = \mathbb{E}[W(t)] + (2p - 1) \)

- Applying recursively \( \Rightarrow \mathbb{E}[W(t+1)] = w_0 + (2p - 1)(t+1) \)
Gambling as LTI system with stochastic input

- Recall the evolution of wealth $W(t + 1) = W(t) + \left(2X(t) - 1\right)$

- View $W(t + 1)$ as output of LTI system with random input $2X(t) - 1$

- Recognize accumulator $\Rightarrow W(t + 1) = w_0 + \sum_{\tau=0}^{t} \left(2X(\tau) - 1\right)$
  - Useful, a lot we can say about sums of random variables

- Filtering random processes in signal processing, communications, ...
Numerical analysis of simulation outcomes

- For a more accurate approximation, analyze simulation outcomes.
- Consider $J$ experiments. Each yields a wealth history $W_j(t)$.
- Can estimate the average outcome via the sample average $\bar{W}_J(t)$

$$\bar{W}_J(t) := \frac{1}{J} \sum_{j=1}^{J} W_j(t)$$

- Do not confuse $\bar{W}_J(t)$ with $\mathbb{E}[W(t)]$.
  - $\bar{W}_J(t)$ is computed from experiments, it is a random quantity in itself.
  - $\mathbb{E}[W(t)]$ is a property of the random variable $W(t)$.
  - We will see later that for large $J$, $\bar{W}_J(t) \to \mathbb{E}[W(t)]$. 

Introduction to Random Processes  
Introduction
Analysis of simulation outcomes: mean

- Expected value $\mathbb{E}[W(t)]$ in black
- Sample average for $J = 10$ (blue), $J = 20$ (red), and $J = 100$ (magenta)
There is more information in the simulation’s output

Estimate the distribution function of \( W(t) \) ⇒ Histogram

Consider a grid of points \( w^{(0)}, \ldots, w^{(M)} \)

Indicator function of the event \( w^{(m)} \leq W_j(t) < w^{(m+1)} \)

\[
\mathbb{1} \left\{ w^{(m)} \leq W_j(t) < w^{(m+1)} \right\} = \begin{cases} 1, & \text{if } w^{(m)} \leq W_j(t) < w^{(m+1)} \\ 0, & \text{otherwise} \end{cases}
\]

Histogram is then defined as

\[
H \left[ t; w^{(m)}, w^{(m+1)} \right] = \frac{1}{j} \sum_{j=1}^{J} \mathbb{1} \left\{ w^{(m)} \leq W_j(t) < w^{(m+1)} \right\}
\]

Fraction of experiments with wealth \( W_j(t) \) between \( w^{(m)} \) and \( w^{(m+1)} \)
Distribution broadens and shifts to the right \( (t = 10, 50, 100, 200) \)
What is this class about?

- Analysis and simulation of **stochastic systems**
  - A system that **evolves in time** with some randomness
- They are usually quite **complex** ⇒ Simulations
- We will learn how to **model** stochastic systems, e.g.,
  - \( X(t) \) Bernoulli with parameter \( p \)
  - \( W(t + 1) = W(t) + 1, \) when \( X(t) = 1 \)
  - \( W(t + 1) = W(t) - 1, \) when \( X(t) = 0 \)
- ... how to **analyze** their properties, e.g., \( \mathbb{E}[W(t)] = w_0 + (2p - 1)t \)
- ... and how to **interpret** simulations and experiments, e.g.,
  - Average tendency through sample average
  - Estimate probability distributions via histograms