Stats 150, Summer 2019
Lecture 1, Thursday, 8/29/2019

CLASS SYLLABUS:
Course Description: Random walks, discrete time Markov chains, Poisson processes. Continuous time Markov chains, queueing theory, point processes, branching processes, renewal theory, stationary processes, Gaussian processes.

Prerequisite: Stat 134 or equivalent.

Text: Essentials of Stochastic Processes by Richard Durrett (available on Springer)

Grading:

$$\text{Overall grade} = \max\{\ 20\%\text{hw} + 30\%\text{mt} + 50\%\text{final},\ 20\%\text{hw} + 80\%\text{final}\ \}$$

Homework is due **weekly by 11:59pm each Thursday** on Gradescope. Lowest 2 homework scores will be dropped. No late homework.

Midterm: In-Class, Thursday Oct 17

Final: Tuesday, Dec 17, 8-11am

Supplemental Sections: Tues 4-6pm : 334 Evans

GSI: Jake Calvert

OH: T/Th 11am-12pm, 2-4pm: 303 Evans

Pitman notes that exams will require no calculations (will be theoretical) and mentions that should we have computational knowledge, he encourages that we practice simulations. We can compute 'the hell out of' problems to numerically get the correct answers, and Pitman encourages this.

Schedule of Topics:

- Week 0, Th Aug 29
  - 1.1 - 1.2 : Markov Chains, Multistep Transitional Prob (p. 1, 9)
- Week 1, T/Th Sept 3, 5
  - 1.3 Classification of States (p. 13)
  - 1.4 Stationary Distributions, Doubly Stochastic Chains (p. 21)
  - 1.5 Detailed Balance Condition, Reversibility, The Metropolis-Hastings Algorithm (p. 28)
- Week 2, T/Th Sept 10, 12
  - 1.6 Limit Behavior (p. 40)
  - 1.7 Returns to a Fixed State (p. 46)
  - 1.8 Proof of the Convergence Theorem (p. 50)
- Week 3, T/Th Sept 17, 19
  - 1.9 Exit Distributions (p. 53)
  - 1.10 Exit Times (p. 61)
- Week 5, T/Th Oct 1, 3
  - 2. Poisson Processes (p. 95)
  - 2.1 Exponential Distribution (p. 95)
  - 2.2 Defining the Poisson Process (p. 100)
  - 2.2.1 Constructing the Poisson Process (p. 103)
  - 2.2.2 More Realistic Models (p. 104)
- Week 6, T/Th Oct 8, 10
  - 2.3 Compound Poisson Processes (p. 106)
  - 2.4 Transformations: Thinning, Superposition, Conditioning (p. 108)
- Week 7, T/Th Oct 15, 17: Review, **Midterm Exam**
- Week 8, T/Th Oct 22, 24
  - 3.1 Renewal Processes, Laws of Large Numbers (p. 125)
  - 3.2 Applications to Queueing Theory (p. 130): GI/G/1 Queue, Cost Equations, M/G/1 Queue
  - 3.3 Age and Residual Life*, Discrete Case + General Case (p. 136)
- Week 9, T/Th Oct 29/31
  - 4.1 Continuous Time Markov Chains (p. 147)
  - 4.2 Computing Transitional Probability: Branching Processes (p. 152)
  - 4.3 Limiting Behavior, Detailed Balance Condition (p. 162)
- Week 10, T/Th Nov 5, 7
  - 4.4 Exit Distributions and Exit Times (p. 170)
  - 4.5 Markovian Queues, Single Server Queues, Multiple Servers, Departure Processes (p. 176)
  - 4.6 Queueing Networks* (p. 183)
- Week 11, T/Th Nov 12, 14
  - 5.1-2 Martingales, Conditional Expectations (p. 201)
  - 5.3-4 Gambling Strategies, Stopping Times, Applications (207)
- Week 12, T/Th Nov 19, 21
  - 5.4.1-2 Applications: Exit Distributions, Exit Times (p. 212)
  - 5.4.3-4 Extinction and Ruin Probabilities, Positive Recurrence of the GI=G=1 Queue* (p. 216)
- Week 13, T/Th Nov 26 : Gaussian Process and Brownian Motion (+ Thanksgiving)
- Week 14, T/Th Dec 3, 5 : Gaussian Processes and Brownian Motion
- Week 15, T/Th Dec 10, 12: RRR Week

1 Course Overview: Stochastic Processes

We may ask: What is a Stochastic Process? We can do a lot worse than go to Wikipedia. Most of our ideas and concepts are very mainstream, and we can simply Google them. The Wikipedia pages are mostly high-quality, according to Pitman. By the time we finish the course, we will likely cover the contents on those Wikipedia pages.

A pretty good start to answering this is to say that a Stochastic Process is a collection of random variables indexed by a parameter set I. That is,

$$(X_i,\ i \in I),$$

and usually we usually take I as **time**, but it could be space (or even fancier, could be both).
Behind these random variables, there is some **probability measure** which we will call $\mathbb{P}$.

We sometimes denote the measure as $\mathbb{P}_\lambda$, and start with the friendly non-negative integers: $I = \{0, 1, 2, 3, \dots\}$.

$$\begin{aligned}
X_0 &= \text{Initial State} \\
X_1 &= \text{State with Time 1} \\
&\vdots \\
X_n &= \cdots
\end{aligned}$$

As the first interesting process, we want to think of the **Markov Chain** with a <u>countable space</u> S and a <u>stationary transitional probability matrix P</u>.

Notice there is a slight difference between Pitman's blackboard notation and that given in the text. When Pitman writes $\mathbb{P}$ as a probability measure on the blackboard, the text will simply write P. Pitman will write P where the text will write p for a **matrix** (although Pitman does not encourage using this lowercase letter to denote a matrix).

Remark: $\mathbb{P}$ denotes a probability measure.
P denotes a matrix.

To specify a Stochastic Process (S.P.), we must decide the joint distribution of its variables. That is, we want to know the probability of :

$$\mathbb{P}(X_0 \le 3, X_1 \le 5, X_2 \le 7) = \text{something}$$

This is a joint probability involving 3 variables X_0, X_1, X_2. The idea behind an S.P. is that this does not matter on how many variables at which we look. We simply use the probability measure $\mathbb{P}$.

A question is posed in class: What is the **sample space** (outcome space) Ω here? The quick answer given by Pitman is that it depends. The canonical answer is:
For state space S and Time set I, the Ω space is:

$$\begin{aligned}
\Omega &= \prod_{i \in I} S_i \\
&= \text{all }\ (x_i,\ i \in I),
\end{aligned}$$

where S_i is a copy of S and $x_i \in S$. If we take an infinite horizon $I = \{0, 1, 2, \dots\}$, then we have:

$$\Omega = \text{all spaces } (x_0, x_1, x_2, \dots), \quad x_i \in S$$

2 Markov Chain

Markov Chains are a bit more interesting, where we have some dependence between variables, so we cannot simply multiply together measures or sample spaces.
How do we define probabilities for a Markov Chain (MC)? Let's assume S is the set of nonnegative integers, $\mathbb{N}_0 = \{0, 1, 2, \dots\}$. We call this a **counting variable**. Then we have:

$$\mathbb{P}\left(X_0 \le 3, X_1 \le 5, X_2 \le 7\right) = \sum_{x=0}^{3} \sum_{y=0}^{5} \sum_{z=0}^{7} P(x, y, z),$$

where P is a matrix and we add up over all the cases. The **Markov Chain** specifies the joint probability $P(x, y, z)$ in a very simple way:
It gives the joint probability:

$$\begin{aligned} P(x, y, z) &= \mathbb{P}(X_0 = x) \cdot \underbrace{P(x, y)}_{\text{TPM}} \cdot P(y, z) \\ &= \mathbb{P}(X_0 = x) P(y|x) = P(z|y) \quad \text{(we do not use this notation)} \end{aligned}$$

where TPM stands for 'transition probability matrix', and

$$P(x, y) = \mathbb{P}\left(X_1 = y \mid X_0 = x\right),$$

where

$$P(AB) = P(A \cap B) = P(A)P(B|A)$$

Notice the inversion of the variables here in our choice not to use $P(z|y)$ notation.

Now if we want to look at $P(x, y, z, w)$, to modify our formula, we take:

$$P(x, y, z, w) = \mathbb{P}(X_0 = x)P(x, y)P(y, z)P(z, w)$$

Remark: Notice that for any sequence of 3 random variables (RV) x_0, x_1, x_2, we can write the probability that we have the triple of values:

$$\begin{aligned} &\mathbb{P}\left(X_0 = x, X_1 = y, X_2 = z\right) \\ &= \mathbb{P}\left(X_0 = x\right) \mathbb{P}\left(X_1 = y | X_0 = x\right) \mathbb{P}\left(X_2 = z | X_0 = x, X_1 = y\right) \end{aligned}$$

What is special about this formula is that the last factor only has y and is **independent of** x. This is the key to the Markov Chain probability:

$$\mathbb{P}\left(X_2 = z | X_0 = x, X_1 = y\right) = \mathbb{P}\left(X_2 = z | \overbrace{X_1 = y}\right)$$

Definition: Markov Property -

Generally $X_0, X_1, X_2, \ldots$ has the **Markov Property** if

$$\begin{aligned}\mathbb{P}\left(X_{n+1} = x_n | X_0 = x_0, X_1 = x_1, \ldots, X_n = x_n\right) \\ = \mathbb{P}\left(X_{n+1} = x_{n+1} | X_n = x_n\right)\end{aligned}$$

Or more simply, as a mantra: Past and future are **conditionally independent** given the present.

2.1 Homogeneous Transition Probabilities

One more point: We are assuming that our Markov Property has a **homogeneous transition probabilities**, where we use the **same** rule (matrix P) to go from one state to the next (as opposed to changing the rule at each step).
Now for any assigned initial distribution λ for X_0,

$$\mathbb{P}(X_0 = x_0) = \lambda(x_0), \qquad \sum_{x_0}^{\lambda(x_0)} = 1, \quad \lambda(x_0) \geq 0,$$

and for any transitional probability P,

$$P(x, y) \geq 0, \qquad \sum_y P(x, y) = 1,$$

which we say are the **rules of a transition matrix**.

From initial λ and a matrix P, we now have:

$$\mathbb{P}\left(\cap_{i=0}^n (X_i = X_i)\right) = \lambda(x_0) \cdot \prod_{i=0}^{n-1} P(x_i, x_{i+1})$$

This is a concise way to say that we make a next move given the current state using the same matrix P. This is the prescription of the joint distribution of the first $n + 1$ steps of a Markov Chain with initial distribution λ and a homogeneous TPM (Transition Probability Matrix) P. It is more or less obvious that this is a proper (rules of probability) assignment of a joint distribution.
To check this, we want to see that summing all outcomes gives 1 and that all probabilities are nonnegative (trivial). That is,

$$\text{Prob} \geq 0, \qquad \text{Probs sum to 1.}$$

To see the second part via some inductive or iterative argument, notice that we have:

$$\sum_{x_0} \sum_{x_1} \cdots \left(\sum_{x_n} \xi \right) = 1$$

due to our previous requirement above (pink) $\sum_y P(x, y) = 1$. If we are pedantic, we may want to formalize via induction; however, for our purposes it suffices to recognize the argument. This is saying that we can certainly make a probability assignment as well as that we can simulate this guy.

2.2 Simulations

We have a supply of independent uniform $[0,1]$ variables: $U_0, U_1, U_2, \dots$. (Of course, these random numbers are not truly random as given by pseudorandom numbers). Any random process we discuss in this course can be simulated by some uniform variables.

First of all, how do we make $X_0 \sim \lambda$? To simplify, let's say $S = \{0, 1, 2, \dots\}$. We can take U_0 to make X_0 (notice that the sample space need not be a sequence space). That is, set $X_0 := 0$ if $U_0 \in [0, \lambda(0)]$, and

$$X_0 := \begin{cases} 0, & U_0 \in [0, \lambda(0)] \\ 1, & U_0 \in (\lambda(0), \lambda(0) + \lambda(1)] \\ 2, & U_0 \in (\lambda(0) + \lambda(1), \lambda(0) + \lambda(1) + \lambda(2)] \\ \vdots, & \vdots \end{cases}$$

If $X_0 = x_0$, then:

$$X_1 = 0, \text{ if } U_1 \in [0, P(x_0, 0)],$$

and

$$X_1 = 1, \text{ if } U_1 \in (P(x_0, 0), P(x_0, 0) + P(x_0, 1)],$$

and so on. If $X_0 = x_0$ and $X_1 = x_1$, then make $X - 2$ from $P(x_1, \cdot)$. This would take about 10 lines of Python or R code, and Pitman says we should really be able to do this.

Example: Gambler's Ruin
Consider a gambler with an initial fortune \$$a$ with $0 \le a \le N$ and $a \in \{1, 2, \dots, N-1\}$.
Take $X_0 = a$, and with each play, gambler gains +\$1 with probability p and loses \$1 with probability q where $p + q = 1$.

Pitman notes this is fairly obviously a Markov Chain. X_n is the capital after n states. λ is degenerate with $\lambda(a) = 1$ and $\lambda(x) = 0$ else. We write:

$$\begin{aligned} P(x, x+1) &= p, \qquad 0 < x < N \\ P(x, x-1) &= q \end{aligned}$$

For mathematical convenience, we construct an **absorbing boundary** to say that the game goes on infinitely if the gambler hits 0. We say:

$$P(0,0) = 1, \quad P(N,N) = 1.$$

For $N = 3$, we can simply fill in the matrix:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ q & 0 & p & 0 \\ 0 & q & 0 & p \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

If we take $p = q = \frac{1}{2}$, then it is not too hard to prove that we will hit one of the boundaries in that we will walk out with 0 or N dollars (as opposed to oscillating forever). We reason in class that this case is **linear in** a. There is only one such function that is linear in a and satisfies our given boundary equations, so we conclude this is simply:

$$\frac{a}{N}.$$

For the case where $p \neq q \neq \frac{1}{2}$, then this is more complicated with writing equations with unknown probabilities.

Lecture ends here.