ORIE 6500 - Applied Stochastic Processes

August 30, 2019

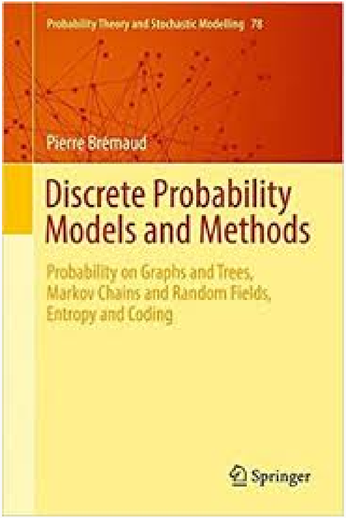
Semester: Fall 2019

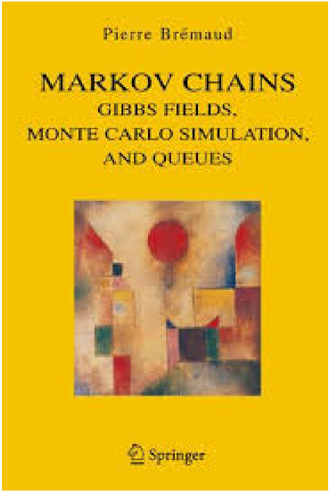
essential course information

- *instructor*: Sid Banerjee, `sbanerjee@cornell.edu`
- *(stand-in lectures for next 2 weeks)*
  Andreea Minca, `acm299@cornell.edu`
- *TA*: Xiaoyang (Andrew) Lu, `xl562@cornell.edu`
- *lectures*: MWF 12:20-1:10pm, Phillips 213
  discussion: M 2:30-4:25pm, Upson 222
- *website*
  `https://piazza.com/cornell/fall2019/orie6500.html`

course textbook

we will use the following books for most topics in the course





the fine print

- *homeworks*:
  10 homeworks, individual submissions, typeset in LaTeX
  due Monday 12pm, on https://cmsx.cs.cornell.edu
- *exams*
  prelim: in recitation, tentatively Oct 21
  final: in regular finals slot
- *grading*: 40% homeworks, 60% exams (either 25%-35% or 0%-60%)

background

- the 'undergraduate' (i.e., classical) view

> *"probability theory is nothing but common sense reduced to calculation"*
>
> – Laplace

background

- the ‘undergraduate’ (i.e., classical) view

> *“probability theory is nothing but common sense reduced to calculation”*
>
> – Laplace

- philosophical problems: Bertrand’s paradox

> “given an equilateral triangle inscribed in a circle, and a random chord, what is the probability the chord is longer than the side of the triangle?

background

- the 'undergraduate' (i.e., classical) view

> *"probability theory is nothing but common sense reduced to calculation"*
>
> – Laplace

- philosophical problems: Bertrand's paradox

> "given an equilateral triangle inscribed in a circle, and a random chord, what is the probability the chord is longer than the side of the triangle?

- mathematical problems: Banach-Tarski paradox

> "a solid ball in 3-dimensional space, can be decomposed into 5 disjoint subsets, which can then be reassembled via translations and rotations to yield two identical copies of the original ball"

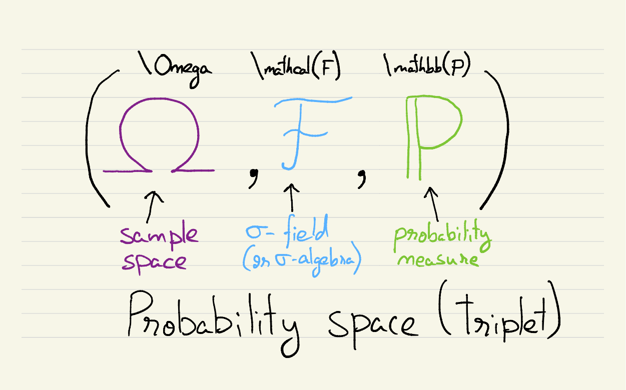
also see `https://www.youtube.com/watch?v=s86-Z-CbaHA`

the solution: measure-theoretic probability

"the theory of probability as mathematical discipline can and should be developed from axioms in the same way as geometry and algebra"

– Andrey Kolmogorov

Kolmogorov's axioms: brief introduction



reading assignment: chapters 1 and 2.1 of Brémaud, Discrete Probability

Kolmogorov's axioms: brief introduction

sample space Ω

$\Omega \triangleq$ set of all possible outcomes ω of any experiment

Kolmogorov's axioms: brief introduction

sample space Ω

$\Omega \triangleq$ set of all possible outcomes ω of any experiment

sigma-field $\mathcal{F}$

collection of subsets of Ω such that:

i. $\Omega \in \mathcal{F}$ (certain event in $\mathcal{F}$)

ii. if $A \in \mathcal{F}$ then $\bar{A} \in \mathcal{F}$ (complements in $\mathcal{F}$)

iii. if $A_1, A_2, \ldots \in \mathcal{F}$ then $\cup_{i=1}^{\infty} A_i \in \mathcal{F}$ (countably-infinite unions in $\mathcal{F}$)

- $\mathcal{F} = \{\Omega, \varnothing\}$
- $\mathcal{F} = \{\Omega, \varnothing, A, \bar{A}\}$ for any $A \subset \Omega$
- for discrete Ω, the power set $\mathcal{F} = 2^{\Omega}$
- for $\omega = \mathbb{R}^n$, the Borel σ-field $\mathcal{B} \equiv$ 'sets with volume'

Kolmogorov's axioms: brief introduction

- $\Omega \triangleq$ all possible outcomes ω of experiment
- $\mathcal{F} \triangleq$ collection of subsets of Ω that includes Ω, complements and countable unions

probability measure $\mathbb{P}$

mapping $\mathbb{P} : \mathcal{F} \to \mathbb{R}$ such that

i. $\mathbb{P}[\Omega] = 1$

ii. $0 \leq \mathbb{P}[A] \leq 1$ for all $A \in \mathcal{F}$

iii. for $A_1, A_2, \ldots \in \mathcal{F}$ such that $A_i \cap A_j = \varnothing \, \forall \, i, j$

$$\mathbb{P}\left[\bigcup_{i=1}^{\infty} A_i\right] = \sum_{i=1}^{\infty} \mathbb{P}[A_i]$$

what is this class about

a 'non measure-theoretic' introduction to 'applied' probability

- overview (without full details) of measure-theoretic concepts – integration, continuity, convergence

what is this class about

a 'non measure-theoretic' introduction to 'applied' probability

- overview (without full details) of measure-theoretic concepts – integration, continuity, convergence
- scaling and asymptotics of discrete probability models (balls in bins, random walks, branching process, probability on networks)

what is this class about

a 'non measure-theoretic' introduction to 'applied' probability

- overview (without full details) of measure-theoretic concepts – integration, continuity, convergence
- scaling and asymptotics of discrete probability models (balls in bins, random walks, branching process, probability on networks)
- the theory of discrete-time Markov chains (classification, convergence, reversibility, mixing)

what is this class about

a ‘non measure-theoretic’ introduction to ‘applied’ probability

- overview (without full details) of measure-theoretic concepts – integration, continuity, convergence
- scaling and asymptotics of discrete probability models (balls in bins, random walks, branching process, probability on networks)
- the theory of discrete-time Markov chains (classification, convergence, reversibility, mixing)
- advanced tools: tail bounds and the probabilistic method, coupling and stochastic ordering, martingales, renewal theory

what is this class about

a ‘non measure-theoretic’ introduction to ‘applied’ probability

- overview (without full details) of measure-theoretic concepts – integration, continuity, convergence
- scaling and asymptotics of discrete probability models (balls in bins, random walks, branching process, probability on networks)
- the theory of discrete-time Markov chains (classification, convergence, reversibility, mixing)
- advanced tools: tail bounds and the probabilistic method, coupling and stochastic ordering, martingales, renewal theory
- intro to continuous stochastic processes: the Poisson point process, continuous-time Markov chains

why should we care?

- beautiful theory with deep connections to measure theory and analysis

why should we care?

- beautiful theory with deep connections to measure theory and analysis
- fundamental way of thinking:

    combinatorics (the probabilistic method)
    economics (game theory)
    physics (statistical physics, quantum mechanics)
    theoretical computer science

why should we care?

- beautiful theory with deep connections to measure theory and analysis
- fundamental way of thinking:

    combinatorics (the probabilistic method)

    economics (game theory)

    physics (statistical physics, quantum mechanics)

    theoretical computer science

- wide range of applications:
    - modeling complex human systems and networks
    - simulation and randomized algorithms
    - working with big data
    - distributed algorithms and cryptography
    - optimization and machine learning

what is ‘scaling’?

credits: www.fixturescloseup.com

warmup example: balls in bins

throw m balls into n bins uniformly at random (u.a.r.)

assume n is very large.

think of number of balls $m(n)$ as a function of n.

warmup example: balls in bins

throw m balls into n bins uniformly at random (u.a.r.)

assume n is very large.

think of number of balls $m(n)$ as a function of n.

three questions

- how big should m be before some bin has at least two balls? (the 'birthday paradox')
- how big should m be before every bin has at least one ball? (the 'coupon-collector problem')
- if $m = n$, how many balls in max-loaded bin?

balls in bins: scaling

throw m balls into n bins uniformly at random (u.a.r.)

three questions

- how big should m be before some bin has at least two balls?
  (the 'birthday paradox')
  answer:
- how big should m be before every bin has at least one ball?
  (the 'coupon-collector problem')
  answer:

  if $m = n$, how many balls are there in the max-loaded bin?
  answer:

balls in bins: scaling

throw m balls into n bins uniformly at random (u.a.r.)

three questions

- how big should m be before some bin has at least two balls?
  (the 'birthday paradox')
  answer: $\Theta\left(\sqrt{n}\right)$
- how big should m be before every bin has at least one ball?
  (the 'coupon-collector problem')
  answer:

  if $m = n$, how many balls are there in the max-loaded bin?
  answer:

balls in bins: scaling

throw m balls into n bins uniformly at random (u.a.r.)

three questions

- how big should m be before some bin has at least two balls?
  (the 'birthday paradox')
  answer: $\Theta\left(\sqrt{n}\right)$
- how big should m be before every bin has at least one ball?
  (the 'coupon-collector problem')
  answer: $\Theta\left(n \log n\right)$

  if $m = n$, how many balls are there in the max-loaded bin?
  answer:

balls in bins: scaling

throw m balls into n bins uniformly at random (u.a.r.)

three questions

- how big should m be before some bin has at least two balls?
  (the 'birthday paradox')
  answer: $\Theta\left(\sqrt{n}\right)$
- how big should m be before every bin has at least one ball?
  (the 'coupon-collector problem')
  answer: $\Theta\left(n \log n\right)$

  if $m = n$, how many balls are there in the max-loaded bin?
  answer: $\Theta\left(\frac{\log n}{\log \log n}\right)$

balls in bins: scaling

throw m balls into n bins uniformly at random (u.a.r.)

three questions

- how big should m be before some bin has at least two balls?
  (the 'birthday paradox')
  answer: $\Theta\left(\sqrt{n}\right)$
- how big should m be before every bin has at least one ball?
  (the 'coupon-collector problem')
  answer: $\Theta\left(n \log n\right)$

  if $m = n$, how many balls are there in the max-loaded bin?
  answer: $\Theta\left(\frac{\log n}{\log \log n}\right)$

takeaway: in large stochastic systems, simple questions have 'interesting' answers

balls in bins: a final twist

throw m balls into n bins uniformly at random (u.a.r.)

> if we choose $m = n$, how many balls are there in the most-loaded bin?
> *answer*: maximum load is $\Theta\left(\frac{\log n}{\log \log n}\right)$

the power of two choices

suppose instead we do the following: for each ball, choose 2 bins u.a.r., and drop ball in less-loaded bin.

throw m balls into n bins uniformly at random (u.a.r.)

if we choose $m = n$, how many balls are there in the most-loaded bin?

answer: maximum load is $\Theta\left(\frac{\log n}{\log \log n}\right)$

the power of two choices

suppose instead we do the following: for each ball, choose 2 bins u.a.r., and drop ball in less-loaded bin.

the maximum load is now $\Theta(\log \log n)$

balls in bins: a final twist

throw m balls into n bins uniformly at random (u.a.r.)

if we choose $m = n$, how many balls are there in the most-loaded bin?

answer: maximum load is $\Theta\left(\frac{\log n}{\log \log n}\right)$

the power of two choices

suppose instead we do the following: for each ball, choose 2 bins u.a.r., and drop ball in less-loaded bin.

the maximum load is now $\Theta(\log \log n)$

takeaway: in large stochastic systems, small changes can lead to dramatically different outcomes