MATH 342: PROBABILITY

INTRO:

- Who am I: Williams, Smith
- Number Theory, Prob, Sabermetrics
- Cam/Kayla

Goals:

- Obviously learn prob
- learn to ask good questions
- learn to model
- Prob + Stats + CS → job (Moneyball Exception)

HAVE FUN!

Grading:

- HW : 20%
- Midterms (1-2) and Final : 30%, 35%
- Project (oral or written) : 10%
- Preparing for class : 5%
  - ↳ Not covering all of book in class, jumping point

Topics:

- Will forget most material if not use, techniques important
- General Probability
- Common distr and Modeling
- Limiting behavior (CLT)

Rest TBD, including Monte Carlo (biggest 20th century paper)

Opportunity to work on book

Assuming comfortable with Appendices

Motivating Examples

Choosing two examples that illustrate key concepts
↳ See also book in progress

Example 1: The Birthday Problem

How many people need in a room to have at least a 50% chance that two share a birthday?

Important to clearly define problem!

- My wife, former suitemate and a thesis advisor all identical twins
- What about Feb 29th?

↳ Must state clearly
↳ assume all days equally likely, no Feb 29th

Do clicker question on Earth and on Pluto

Soln: Brute Force: go through all ways at least 2 share
Nightmare! Exactly 2 on Jan 1, or 3, or 4, or 5...
Double count if 2 on Jan 1 and 2 on June 6

Prob(A happens) = 1 − Prob(A doesn't happen)

Prob(n have distinct bdays): $\frac{365}{365} \cdot \frac{365-1}{365} \cdots \frac{365-(n-1)}{365}$

$$= \prod_{k=0}^{n-1}\left(1-\frac{k}{365}\right) = \frac{365!}{365^n(365-(n-1))!}$$

Try Extreme Cases to get ~~a~~ sense. Clearly $n \le 365$.
If $n = 185$... (too high: final factors exceed $1/2$.

Example 1: Birthday Problem (cont)

- Just b/c have formula doesn't mean "tractable"

  Ans is $\prod_{k=0}^{n-1}\left(1-\frac{k}{365}\right)$, or if N days/year is $\prod_{k=0}^{n-1}\left(1-\frac{k}{N}\right)$.

  How to get to 50%? What is n_{50}? or $n_{50}(N)$?

$$\boxed{\text{FREQUENTLY GO FOR APPROXIMATIONS}}$$

- More tools, more can do
- Pavlovian response:

$$\boxed{\text{See a product, Think of logarithms}}$$

Want $\frac{1}{2} = \prod_{k=0}^{n-1}\left(1-\frac{k}{N}\right)$. Expect n "small" relative to N,

so $\log(1/2) = \sum_{k=0}^{n-1} \log(1-k/N)$. Use $\log(1-x) = -x + O(x^2)$

$\hookrightarrow \log(1-x) = \sum_{\ell=1}^{\infty} -\frac{x^\ell}{\ell}$

so $-\log 2 \approx -\sum_{k=0}^{n-1}\left(\frac{k}{N} + O\left(\frac{k^2}{N^2}\right)\right)$

so $\log 2 \approx \frac{1}{N}\sum_{k=0}^{n-1} k + O\left(\frac{n^3}{N^2}\right)$

$\hookrightarrow$ use either $\int_0^{n-1} x\,dx$ or $\frac{(n-1)(n)}{2}$, both $\approx \frac{n^2}{2}$

or a bit better $\frac{(n-1)n}{2} \approx \frac{(n-1/2)^2}{2}$

so $\log 2 \approx \frac{(n-1/2)^2}{2N} + O\left(\frac{n^3}{N^2}\right)$

$\hookrightarrow$ see $n \approx \sqrt{N}$ so error negligible

so $(n-\frac{1}{2})^2 \approx 2N\log 2$ or $n \approx \sqrt{N\log 4} + \frac{1}{2}$

$\hookrightarrow$ Key number: $\log 2 \approx .69$, so if $N = 365$ estimate is ≈ 23 (see back: Newton's Method for $\sqrt{5}$)

Question: do you think we're over or under estimating with this method?

EXAMPLE 2: ROULETTE

Simplify game: only bet red or black, each equally likely

Double plus One strategy: eventually win \$1!

↳ What are problems with this?

- Unlimited reserves: rich uncle hypothesis
- house limit: severe!

Question: what is prob at least 5 consec black ~~in 100 tosses~~?

Very hard to enumerate possibilities: danger double counting

↳ double counting one of biggest headaches in subject

$\boxed{\text{Key technique: look at Prob(A happens)} = 1 - \text{Prob(A doesn't happen)}}$

Before solving, let's try and get a feel

$\boxed{\text{Power of quick, simple estimates}}$

I find confusing: upper bound for P(no A) is lower bound for Prob(A)

↳ Simple upper bound for Prob(no 5 in 100 tosses)

$\underbrace{\quad\quad}_{5}\ \underbrace{\quad\quad}_{5}\ \cdots\ \underbrace{\quad\quad}_{5}$ (20 times)

$31/32$ or $1 - \frac{1}{32}$ or $1 - \frac{1}{2^5}$ probability don't have BBBBB

$\boxed{\text{Prob of many indep events is product of probabilities}}$

$\text{Prob(no 5 in 100)} \leq \left(1 - \frac{1}{2^5}\right)^{20}$ ($\leq$ as ignores overlap)

So $\text{Prob(5 consec B in 100)} \geq 1 - \left(1 - \frac{1}{2^5}\right)^{20} \approx$

(see back for quick estimate of 47%)

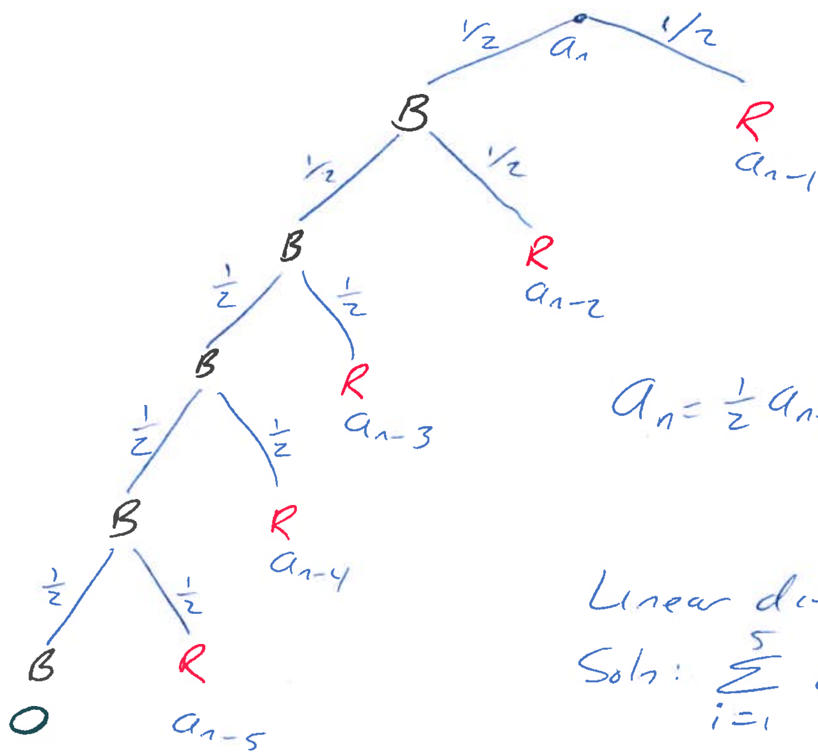
Example 2: Roulette

$\boxed{\text{Technique: Find tractable mathematical model}}$

Let a_n = prob no 5 consec B in n spins; b_n = prob have 5 consec B in n spins

↳ Want b_n, easier find a_n

↳ Small n easy: $a_0 = a_1 = a_2 = a_3 = a_4 = 0$!



$$a_n = \frac{1}{2} a_{n-1} + \frac{1}{4} a_{n-2} + \frac{1}{8} a_{n-3} + \frac{1}{18} a_{n-4} + \frac{1}{32} a_{n-5} + 0$$

Linear difference eq

Soln: $\sum_{i=1}^{5} c_i r_i^n$ where r_i solves

$$r^5 - \frac{1}{2} r^4 - \frac{1}{4} r^3 - \frac{1}{8} r^2 - \frac{1}{16} r - \frac{1}{32} = 0$$

↳ Characteristic polynomial

Note generalizes Fibonacci: $F_{n+1} = F_n + F_{n-1}$

Schilling Paper: $E[\#$ longest run of consec heads$] \approx \log_2 n$, variance is about 4

↳ amazing, indep of n!

Homework ① if have 100 days in year, calc 50% point

② generalize 100 → 5·20 to do blocks of 6: I can do 10·10

⊕ too much trouble

③ use Schilling result to estimate how long can play before VERY likely have at least one run 5 consec

CHAPTER 1: INTRODUCTION

Key Notions:

- Outcome / Sample Space: set of all possible outcomes: Ω
- Events: subsets of Ω, denote often by A or ω
- Probability: fn from subsets of $\mathcal{P}(\Omega) \to [0,1]$

Want prob fn to have "intuitive" properties

↳ $Prob(A \cup B) = Prob(A) + Prob(B)$ if $A \cap B = \emptyset$ (disjoint)

More generally $Prob(\bigcup A_\alpha) = \sum Prob(A_\alpha)$ if pairwise disjoint

↳ Not possible!

Choose X in $[0,1]$ uniformly: all equally likely

↳ What is $Prob(\mathbb{X} = x)$? Indep of x.

$$[0,1] = \bigcup_{x \in [0,1]} \{x\}$$

$$\text{so } 1 = \sum_{x \in [0,1]} Prob(\mathbb{X} = x) = \begin{cases} 0 & \text{if } Prob(\mathbb{X} = x) = 0 \\ \infty & \text{if } Prob(\mathbb{X} = x) > 0 \end{cases}$$

Problem is uncountable sum

Cannot assign probabilities to all subsets of Ω!

Frequently restrict to open / closed intervals and unions and intersections of these (countable), and unions and intersections of those, ...

↳ technically Borel σ-algebra.

CHAPTER 1: SECTION 1.1: EQUALLY LIKELY OUTCOMES

EQUALLY LIKELY OUTCOMES: Ω finite, all elements in Ω equally likely to happen, $Prob(A) = \frac{\#A}{\#\Omega}$.

Examples: rolling die, tossing coins, ...

Skip odds / gambling

Homework: Page 9: #3, #5, #7.

CHAPTER 1.2: INTERPRETATIONS

Frequency interpretation

Probability of A is the expected/estimated relative freq of A in a large number of trials

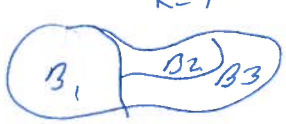
↳ HUGE APPLICATION: MONTE-CARLO INTEGRATION

SECTION 1.3: DISTRIBUTIONS

See table on pg 19

↳ A or B is $A \cup B$, A and B is $A \cap B$

Partition: $B = \bigcup_{k=1}^{n} B_k$ and $B_k \cap B_l = \phi$ if $k \neq l$ (disjoint)



Section 1.3: Distributions (cont)

Rules of Proportion and Probability

- non-neg: $Prob(B) \geq 0$
- Addition: $B_1, \dots, B_n$ partition of $B \Rightarrow Prob(B) = \sum_{k=1}^{n} Prob(B_k)$
- Total one: $Prob(\Omega) = 1$

Lemma: Complement Rule

A^c is not A: $Prob(A^c) = 1 - Prob(A)$

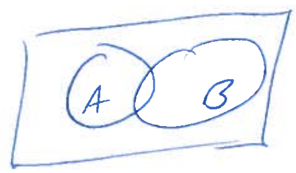
Proof: $A \cup A^c = \Omega$ partition

Lemma: Difference Rule

$A \subset B \to Prob(A) \leq Prob(B)$

$A \subset B \to Prob(B) = Prob(A) + Prob(B \cap A^c)$ (partition)

Lemma: Inclusion/Exclusion



$Prob(A \cup B) = Prob(A) + Prob(B) - Prob(A \cap B)$

More generally:

$$Prob(A_1 \cup \cdots \cup A_n) = \sum_{i=1}^{n} Prob(A_i) - \sum_{1 \leq i < j \leq n} Prob(A_i \cap A_j) + \sum_{1 \leq i < j < k \leq n} Prob(A_i \cap A_j \cap A_k) - \cdots$$

Very powerful: A_i's not need disjoint

Proof: Say x in k of the A_i's. Get $0 = 0$ if $k = 0$,

else get $1 \stackrel{?}{=} \binom{k}{1} - \binom{k}{2} + \binom{k}{3} - \cdots \pm \binom{k}{k}$

subtract: $\binom{k}{0} - \binom{k}{1} + \binom{k}{2} + \cdots \pm \binom{k}{k} = (1-1)^k = 0$

Application: $n!$ permutations, estimate as $n \to \infty$ the prob leave at least one element alone. A_i = event i not moved...

Section 1.3: Distributions (cont)

Bernoulli:

Prob p have success (1) and $1-p$ have failure (0)

Binomial:

Sum of Bernoulli: do n times (indep)

↳ Prob k of n success is $\binom{n}{k} p^k (1-p)^{n-k}$

Uniform:

↳ if set finite, each element taken @ prob $1/\#\Omega$

↳ on interval $[a,b]$: Prob in $[c,d] \subset [a,b]$ is $\frac{d-c}{b-a}$

↳ Can generalize to other shapes

Homework: Page 30: #3, #10, #13; Prove $A \subset B \Rightarrow Prob(A) \leq Prob(B)$

Extra Credit: #16a

SECTION 1.4: CONDITIONAL PROB AND INDEP

Lots of great examples to read

Conditional probability: knowing B happens can influence prob A happens

↳ Ex: fair coin: toss at least 2 heads: event A: prob is 1/2
first toss is a ~~tail~~ head: 1/2 (event B)
$P(A|B)$ is prob of A given B; here it's $3/4 > 1/2$

Formula for P(A|B)

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad (\text{assuming } P(B) \neq 0)$$

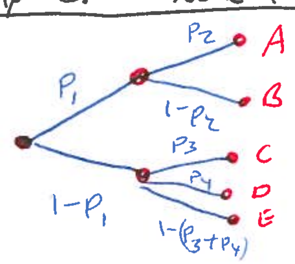
Coming from $\frac{\#(A \cap B)}{\# B} =$

↳ Ex: two sided card: really $\underline{\underline{3}}$ sides that could see, $\underline{\underline{\text{not } 2}}$ cards

EQUIVALENTLY, MULTIPLICATION RULE

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

Multiplication Rule in Tree Diagram



Probability of any is product of the probabilities

using tree, get $P(A) = P(A|B)\,P(B) + P(A|B^c)\,P(B^c)$

Section 1.4: Cond Prob + Indep

Independence (2 Events)

Prob A happens is indep of B happening

$P(A|B) = P(A|B^c) = p_A \implies P(A) = P(A|B)P(B^c) + P(A|B^c)P(B^c) = p_A$

As $P(A \cap B) = P(A|B)P(B)$,

Mult Rule for Indep Events

A, B indep $\implies P(A \cap B) = P(A)P(B)$

$\boxed{\text{One of most important ideas in probability}}$

Question: if A and B, B and C, A and C indep, are all three indep? IE, must $P(A \cap B \cap C) = P(A)P(B)P(C)$?

Homework: #1, #3, #9, Prove or disprove indep question
Page 45

Section 1.5: Bayes' Rule

Extremely important in statistics

Will only do box on page 49, Example 3 on pg 50

Bayes' Rule

Partition $B_1, \ldots, B_n$ of Ω, then

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{P(A|B_1)P(B_1) + \cdots + P(A|B_n)P(B_n)}$$

Proof: $P(A \cap B_i) = P(A|B_i)P(B_i) = P(B_i|A)P(A)$

Use $P(A) = P(A|B_1)P(B_1) + \cdots + P(A|B_n)P(B_n)$

Great Example

test: if have disease, get positive result 95% of time

if don't have disease, get " " 2% " " (false positive

Assume 1% of population has disease

Question: if test is pos, what is chance have disease?

$$P(D|+) = \frac{P(+|D)P(D)}{P(+|D)P(D) + P(+|D^c)P(D^c)}$$

$$= \frac{(.95)(.01)}{(.95)(.01) + (.02)(.99)} = \frac{95}{293} \approx \underline{\underline{32\%}}$$

Section 1.6: Sequences of Events

Only doing some of chapter

Geometric Distribution (Example 2, Pg 58)

Keep rolling till heads

↳ prob p of head, then Prob(1st H after n tosses) $= (1-p)^{n-1}p$

- Reconsider birthdays problem

Independence:

Have n events, need 2^n conditions (one for each path)

Ex: A = roll 7 on two fair die

B = 1st roll is even

C = 2nd roll is even

Homework: Pg 71: #5, #6