Updated 9/10/2024

Part 1: Theory
L02: Basics of Probability (1/2)
[Random experiment, independence, conditional probability, chain rule, Bayes' theorem, random variables]

Javed Aslam, Wolfgang Gatterbauer

cs7840 Foundations and Applications of Information Theory (fa24)

https://northeastern-datalab.github.io/cs7840/fa24/

9/9/2024

Pre-class conversations

- Last class recapitulation
- We strive to keep it interactive, also among us faculty
- Why class slides need up to 2 days after a class
- Office hours: Usually right after class, or via email / Teams
- Organizational matters: Piazza messages? Canvas pictures did not display correctly?
- New class arrivals

- Today:
  - The basics of probability theory

Last time

- Introduction
- Course logistics
- Entropy
  - examples

Today

- Probability primer

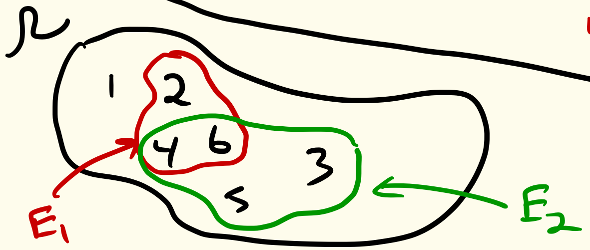
Next time

- Basic concepts of Info. Theory

Probability

- Random experiment
- generate **outcomes** $\omega \in \Omega$
- **sample space**: set of **all possible** outcomes Ω
- **Event**: subset of sample space
- $p: \Omega \to \mathbb{R}$ **probability measure**
  - $0 \le p(\omega) \le 1 \quad \forall \omega \in \Omega$
  - $\sum_{\omega \in \Omega} p(\omega) = 1$

we will assume for now that $p(\omega) = \frac{1}{|\Omega|} \; \forall \omega \in \Omega$



Example

- roll a fair six-sided die
- roll a 5
- $\{1, 2, 3, 4, 5, 6\} = \Omega$
- $E_1 = \text{"even"} = \{2, 4, 6\}$
  $E_2 = \text{"}\ge 3\text{"} = \{3, 4, 5, 6\}$
- $p(1) = p(2) = p(3) = \cdots = p(6) = 1/6$

$$P(E) = \sum_{\omega \in E} p(\omega)$$

e.g. $P(E_1) = p(2) + p(4) + p(6) = 1/6 + 1/6 + 1/6 = 1/2$

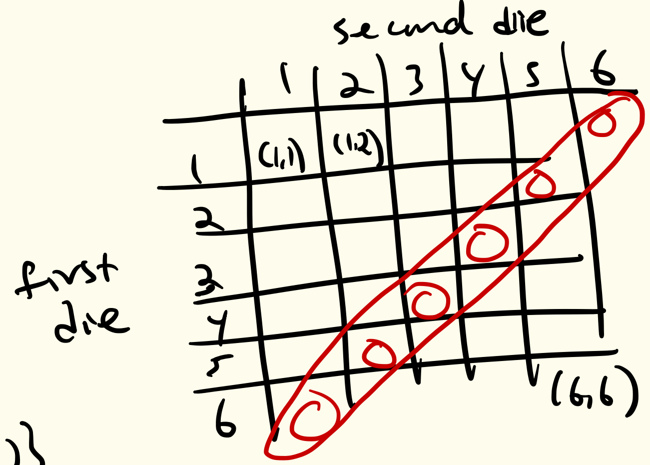
If $p(\omega) = \frac{1}{|\Omega|} \; \forall \omega \in \Omega$

$P(E) = |E| / |\Omega|$ $\qquad P(E_1) = \frac{|E|}{|\Omega|} = \frac{3}{6}$

Examples

① Roll one fair die

$\Omega = \{1,2,3,4,5,6\}$



② Roll <u>two</u> fair die

$$\Omega = \{(1,1),(1,2),(1,3),\ldots(2,1),(2,2),\ldots(6,6)\}$$

$$= \{1,2,3\ldots6\} \times \{1,2,\ldots,6\}$$

$$E_1 = \text{total is } 7 = \{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}$$

$$P(E_1) = \frac{|E_1|}{|\Omega|} = \frac{6}{36} = 1/6$$

$$E_2 = \text{total is greater than } 8$$

$$= 9 \text{ or } 10 \text{ or } 11 \text{ or } 12$$

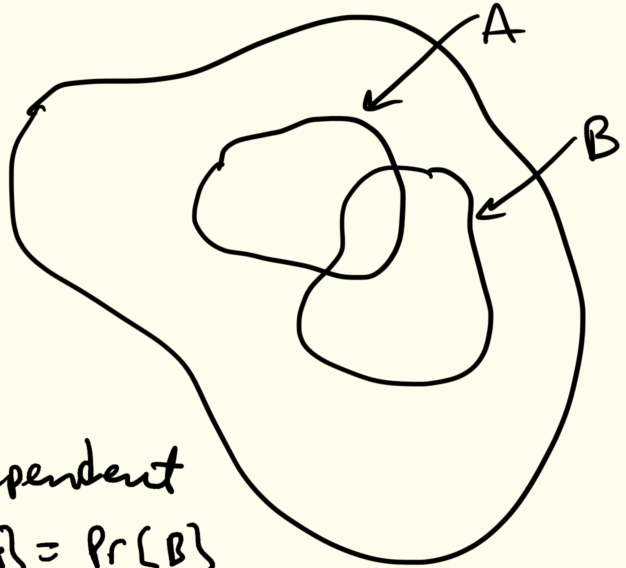
$$|E_2| = \underset{\uparrow 12}{1} + \underset{\uparrow 11}{2} + \underset{\uparrow 10}{3} + \underset{\uparrow 9}{4}$$

$$= 10$$

$$P(E_2) = \frac{|E_2|}{|\Omega|} = \frac{10}{36} = \frac{5}{18}$$

① $\Pr\{A|B\} = \dfrac{\Pr\{A \cap B\}}{\Pr\{B\}}$

② $\Pr\{B|A\} = \dfrac{\Pr\{A \cap B\}}{\Pr\{A\}}$



Independence: Two events A & B are independent if $\Pr\{A|B\} = \Pr\{A\}$, $\Pr\{B|A\} = \Pr\{B\}$

$\Rightarrow$ Knowing B does not change your belief in A.
Knowing A does not change your belief in B.

Claim: A & B are independent if and only if $\Pr\{A \cap B\} = \Pr\{A\} \cdot \Pr\{B\}$

① $\Pr\{A|B\} = \dfrac{\Pr\{A\cap B\}}{\Pr\{B\}}$

Two events A & B are independent if $\Pr\{A|B\} = \Pr\{A\}$, $\Pr\{B|A\} = \Pr\{B\}$

② $\Pr\{B|A\} = \dfrac{\Pr\{A\cap B\}}{\Pr\{A\}}$

Claim: A & B are independent if and only if $\Pr\{A\cap B\} = \Pr\{A\}\cdot\Pr\{B\}$

Proof:

($\Rightarrow$) If $\Pr\{A|B\} = \Pr\{A\}$, then by ① $\Pr\{A\cap B\} = \Pr\{A|B\}\cdot\Pr\{B\} = \Pr\{A\}\cdot\Pr\{B\}$

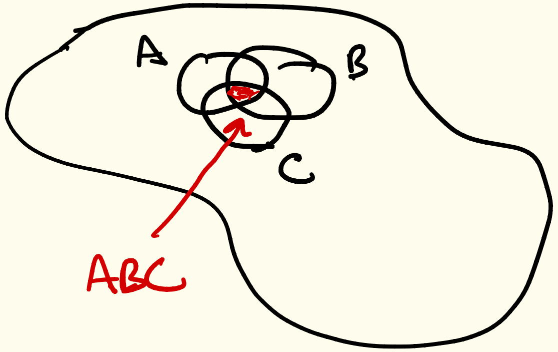
($\Leftarrow$) If $\Pr\{A\cap B\} = \Pr\{A\}\cdot\Pr\{B\}$, then by ① & ②

$$\Pr\{A|B\} = \frac{\Pr\{A\cap B\}}{\Pr\{B\}} = \frac{\Pr\{A\}\cdot\cancel{\Pr\{B\}}}{\cancel{\Pr\{B\}}} = \Pr\{A\}$$

$$\Pr\{B|A\} = \frac{\Pr\{A\cap B\}}{\Pr\{A\}} = \frac{\cancel{\Pr\{A\}}\cdot\Pr\{B\}}{\cancel{\Pr\{A\}}} = \Pr\{B\} \qquad \therefore$$

Chain Rules

- $\Pr\{ABC\} = ?$
  - treat AB as an event



$$\Pr\{ABC\} = \Pr\{AB\} \cdot \Pr\{C \mid AB\}$$

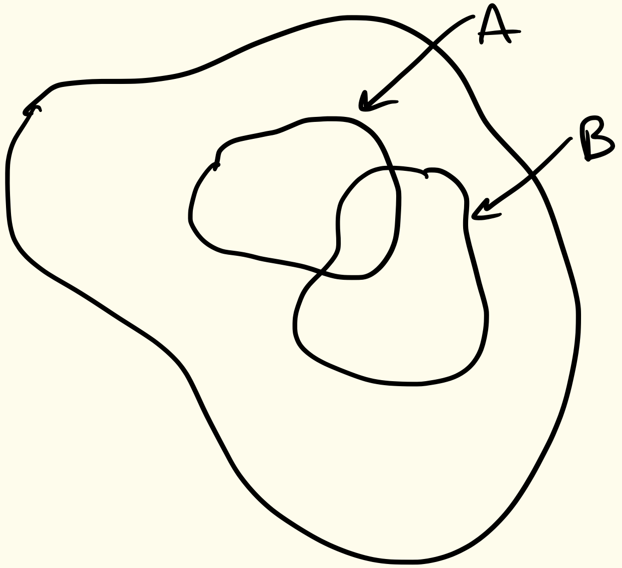
$$= \Pr\{A\} \cdot \Pr\{B \mid A\} \cdot \Pr\{C \mid AB\}$$

Generally: $\Pr\{X_1 X_2 \cdots X_n\} = \Pr\{X_1\} \cdot \Pr\{X_2 \mid X_1\} \cdot \Pr\{X_3 \mid X_1 X_2\} \cdots \Pr\{X_n \mid X_1 \cdots X_{n-1}\}$

Bayes Law

① $\Pr\{A|B\} = \dfrac{\Pr\{A \cap B\}}{\Pr\{B\}}$

② $\Pr\{B|A\} = \dfrac{\Pr\{A \cap B\}}{\Pr\{A\}}$

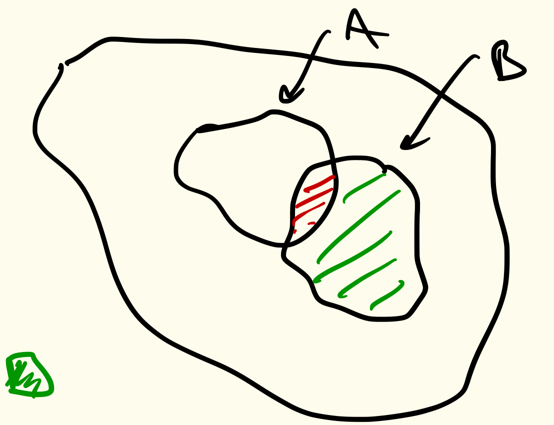


$$\underbrace{\Pr\{A|B\} \cdot \Pr\{B\} = \Pr\{A \cap B\}}_{①} = \Pr\{B|A\} \cdot \Pr\{A\}$$

Solving for $\Pr\{A|B\}$ ...

$$\boxed{\Pr\{A|B\} = \frac{\Pr\{B|A\} \cdot \Pr\{A\}}{\Pr\{B\}}}$$ Bayes Law

- $$\Pr\{A \mid B\} = \frac{\Pr\{B \mid A\} \cdot \Pr\{A\}}{\Pr\{B\}}$$



- what is $\Pr\{B\}$?

$$\Pr\{B\} = \boxed{} + \boxed{}$$

$$= \Pr\{A \wedge B\} + \Pr\{\bar{A} \wedge B\}$$

$$= \Pr\{B \mid A\} \cdot \Pr\{A\} + \Pr\{B \mid \bar{A}\} \cdot \Pr\{\bar{A}\}$$

$$\Pr\{A \mid B\} = \frac{\Pr\{B \mid A\} \cdot \Pr\{A\}}{\Pr\{B \mid A\} \cdot \Pr\{A\} + \Pr\{B \mid \bar{A}\} \cdot \Pr\{\bar{A}\}}$$

$$\Pr\{H \mid E\} = \frac{\Pr\{E \mid H\} \cdot \Pr\{H\}}{\Pr\{E \mid H\} \cdot \Pr\{H\} + \Pr\{E \mid \bar{H}\} \cdot \Pr\{\bar{H}\}}$$

H: hypothesis — patient has Zika

E: evidence — patient tested positive on Zika blood test

Example: Zika in FL in 2016

- prevalence of Zika in FL $\Pr(\text{Zika}) = 10^{-5}$ (1 in 100,000)

- accuracy of blood test is 99%

false negative rate

e.g. $\Pr\{\text{pos. test} \mid \text{Zika}\} = 0.99$ or $\Pr\{\text{neg. test} \mid \text{Zika}\} = 0.01$

$\Pr\{\text{pos. test} \mid \text{no Zika}\} = 0.01$ $\Pr(\text{neg. test} \mid \text{no Zika}) = 0.99$

false positive rate

- patient tests <u>positive</u>: what is chance they have Zika?

$\Longrightarrow$ <u>Not</u> $\Pr\{\text{pos test} \mid \text{Zika}\} = 0.99$

$\Longrightarrow$ You want $\Pr\{\text{Zika} \mid \text{pos. test}\}$

$$\Pr\{H \mid E\} = \frac{\Pr\{E \mid H\} \cdot \Pr\{H\}}{\Pr\{E \mid H\} \cdot \Pr\{H\} + \Pr\{E \mid \overline{H}\} \cdot \Pr\{\overline{H}\}}$$

$$\Pr\{\text{zika} \mid \text{pos test}\} = \frac{\Pr\{\text{pos. test} \mid \text{zika}\} \cdot \Pr\{\text{zika}\}}{\Pr\{\text{pos. test} \mid \text{zika}\} \cdot \Pr\{\text{zika}\} + \Pr\{\text{pos. test} \mid \text{not zika}\} \cdot \Pr\{\text{not zika}\}}$$

$$= \frac{0.99 \cdot 10^{-5}}{0.99 \cdot 10^{-5} + 0.01 \cdot (1 - 10^{-5})}$$

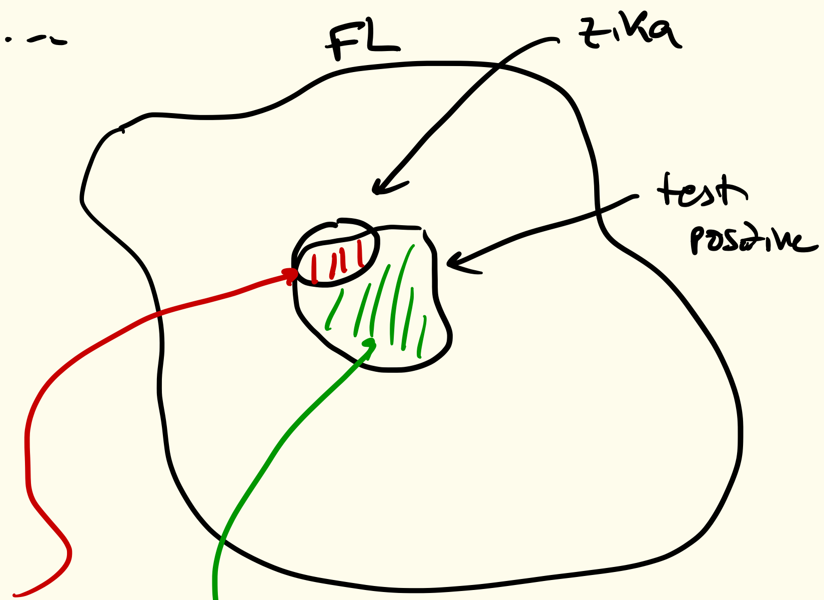
$$= \frac{0.0000099}{0.0000099 + 0.0099999}$$

$$\approx 0.00099$$

$$\approx 0.1\%$$ i.e. about 1 in 1,000!

Seems wildly counterintuitive, but...

- $10{,}000{,}000$ people in FL $\quad 10^7$
- w/ zika? $\quad 10^7 \cdot 10^{-5} = 10^2 = 100$
- w/o zika $\quad 10^7 - 10^2 = 9{,}999{,}900$



---

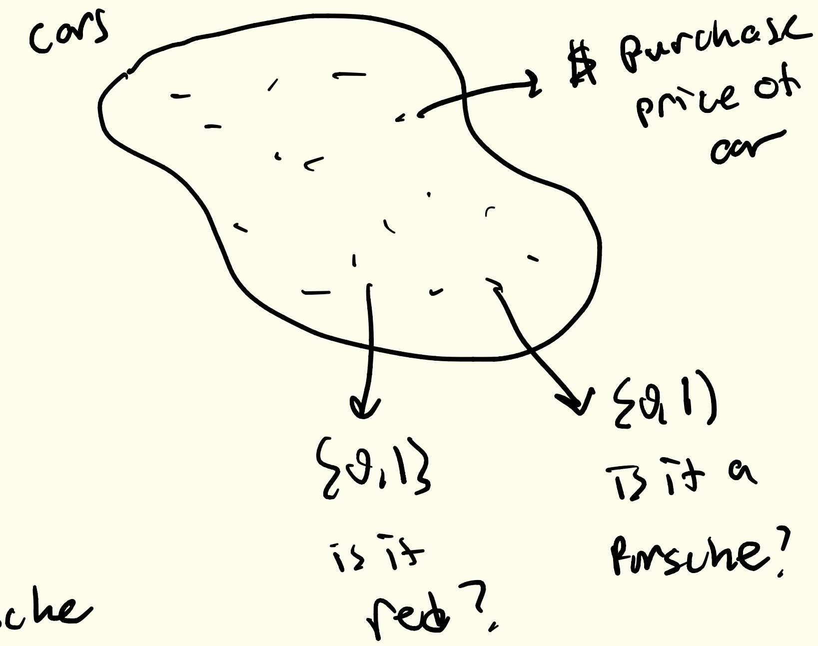
test pos w/ zika: $\quad 100 \cdot 0.99 = 99$

test pos. w/o zika: $\quad 9{,}999{,}900 \cdot 0.01 = 99{,}999$

$\therefore$ among those who test pos,
only 99 out of $(99 + 99{,}999)$
actually have zika – about 1 in 1000.

Random Variable

$X : \Omega \to \mathbb{R}$



Two primary types:

① "numerical"
e.g. $X(\omega) =$ price of ω (dollars)

② "indicator"
e.g. $X(\omega) = \begin{cases} 1 & \text{if } \omega \text{ is a Porsche} \\ 0 & \text{if not} \end{cases}$

$\Rightarrow$ X is a "random" variable because it depends on the outcome of a <u>random</u> <u>experiment</u>

$\Rightarrow$ Underlying <u>probability</u> <u>measure</u> $p : \Omega \to \mathbb{R}$ <u>induces</u> a <u>distribution</u> D over the range of the random variable

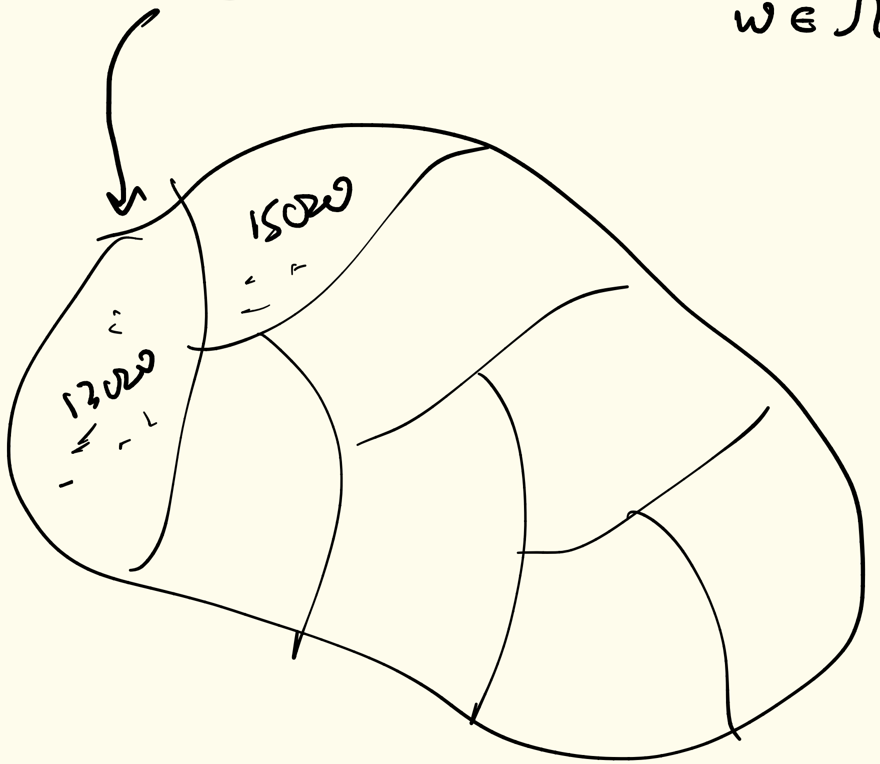
$D : \mathbb{R} \to \mathbb{R}$

Expectation

- Random variable

  $X : \Omega \to \mathbb{R}$

- $E[X] = \sum_{x} x \cdot \Pr[X=x]$ $\qquad E[X] = \sum_{\omega \in \Omega} X(\omega) \cdot p(\omega)$



Updated 9/13/2024

Part 1: Theory
L03: Basics of Probability (2/2)
[Expectation, Variance, Markov chains]

Javed Aslam, Wolfgang Gatterbauer

cs7840 Foundations and Applications of Information Theory (fa24)

https://northeastern-datalab.github.io/cs7840/fa24/

9/11/2024

Pre-class conversations

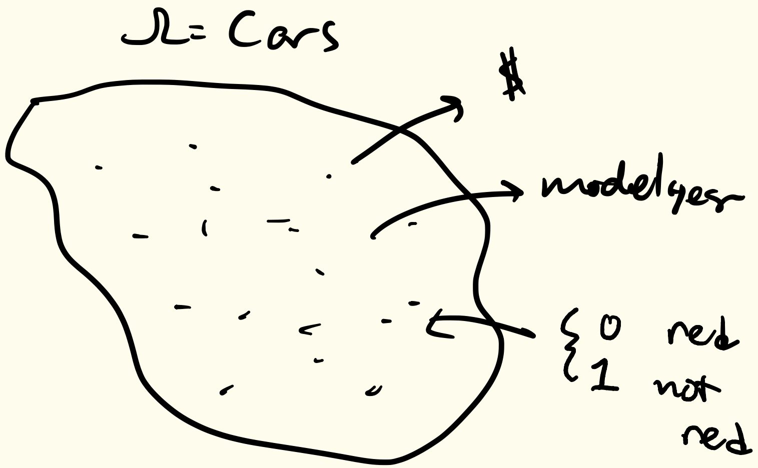
- Last class recapitulation
- PDF Class slides organized by topic (with subsections for classes)
- Any organizational matters: Piazza messages? Organizational matters?
- New class arrivals?

- Today:
  - The basics of probability theory
  - Intuition behind "information" (and "information measures")

Expectation

- Random variable

$X : \Omega \to \mathbb{R}$



$$E[X] = \sum_x x \cdot \Pr[X=x]$$

$$E[X] = \sum_{\omega \in \Omega} X(\omega) \cdot p(\omega)$$

expected or "average" value

1/2 cars	\$27,000
1/3 cars	\$10,000
1/6 cars	\$15,000

$$E[X] = \sum_x x \cdot \Pr[X=x] = \tfrac{1}{2} \cdot 2700 + \tfrac{1}{3} \cdot 10{,}000 + \tfrac{1}{6} \cdot 1500 = \$19{,}333.33$$

Exaple

- Pay \$6 to play game
- Roll two fair 6-sided die
- Pay you sum of die faces

  <u>except</u> if doubles, then 0
- X = winnings (profit)

die 2

die1	1	2	3	4	5	6
1	-6	-3	-2	-1	0	1
2	-3	-6	-1	0	1	2
3	-2	-1	-6	1	2	3
4	-1	0	1	-6	3	4
5	0	1	2	3	-6	5
6	1	2	3	4	5	-6

- $E[X] = \sum_x x \cdot \Pr\{X=x\}$ ← induced distribution

$x=-6 \quad \Pr\{X=-6\} = 6/36 = 1/6$

$x=-3 \quad \Pr\{X=-3\} = 2/36 = 1/18$

$\vdots$

$x=+5 \quad \Pr\{X=5\} = 2/36 = 1/18$

$$E[X] = \sum_x x \cdot \Pr\{X=x\}$$
$$= (-6)\cdot 1/6 + (-3)\cdot 1/18 + \cdots + 5(1/18) = -0.1\overline{66}$$

- $E[X] = \sum_{\omega \in \Omega} X(\omega) \cdot p(\omega)$

$$= \frac{1}{36} \sum_{\omega \in \Omega} X(\omega)$$

$$= \frac{1}{36}\left(\text{"sum whole table"}\right)$$

$$= -6/36 = -0.1\overline{66}$$

$\Rightarrow$ lose 16¢ per play, on average

More on Expectation

Example:
- Roll two fair 6-sided die
- Let X = sum of die faces
- Q: $E[X]$?

die 2

die 1	1	2	3	4	5	6
1	(1,1)	(1,2)				
2	(2,1)			(2,4)		
3						
4						
5						
6						(6,6)

$X : \Omega \to \mathbb{R}$

e.g. $X((2,4)) \to 6$

$$E[X] = \sum_x x \cdot \Pr[X=x]$$

$X=2 \quad \Pr[X=2] = 1/36$

$X=3 \quad \Pr[X=3] = 2/36$

$X=4 \quad \Pr[X=4] = 3/36$

$\vdots$

$X=12 \quad \Pr[X=12] = 1/36$

$$E[X] = \sum_x x \cdot \Pr[X=x]$$

$$= 2 \cdot 1/36 + 3 \cdot 2/36 + \dots + 12 \cdot 1/36 = 7$$

$$E[X] = \sum_{\omega \in \Omega} X(\omega) \cdot p(\omega)$$

$$= \sum_{\omega \in \Omega} X(\omega) \cdot 1/36$$

$$= 1/36 \cdot \sum_{\omega \in \Omega} X(\omega)$$

$$= 1/36 \cdot [\text{sum of table}]$$

$\vdots$

$$= 7$$

Linearity of Expectation

Let X_1 = r.v. for first die roll

Let X_2 = r.v. for second die roll

Let $X = X_1 + X_2$

$$E[X] = E[X_1 + X_2] \overset{*}{=} E[X_1] + E[X_2] = \underbrace{3.5 + 3.5}_{7}$$

Variance & standard deviation

Case 1	case 2	case 3
4'10" 5' 5'2"	4' 5' 6'	3' 5' 7'
$E[X_1] = 5'$	$E[X_2] = 5'$	$E[X_3] = 5'$

How to measure "variability"

3 ways

① $Y_1 = X - E[X]$ × neg & pos cancel

② $Y_2 = |X - E[X]|$ - mean absolute deviation

③ $Y_3 = (X - E[X])^2$ - variance

case 2

① $E[Y_1] = \sum_{\omega \in \Omega} Y(\omega) \cdot p(\omega)$

$= -12'' \cdot 1/3 + 0'' \cdot 1/3 + (+12'') \cdot 1/3$

$= 0''$

② $E[Y_2] = |-12''| \cdot 1/3 + |0''| \cdot 1/3 + |12''| \cdot 1/3$

$= 12 \cdot 1/3 + 0 \cdot 1/3 + 12 \cdot 1/3 = 8''$

③ $E[Y_3] = (-12'')^2 \cdot 1/3 + (0'')^2 \cdot 1/3 + (12'')^2 \cdot 1/3$

$= 96 \text{ in}^2$

$\Rightarrow$ take square root, get standard deviation $\sqrt{96 \text{ in}^2} = 9.8$ in

$$\sigma^2 = Var(X) = E\{(X - E\{X\})^2\}$$ – <u>variance</u>

$$\Rightarrow \quad \sigma = \sqrt{Var(X)} = \sqrt{E\{(X-E\{X\})^2\}}$$ – <u>standard deviation</u>

– back in <u>original</u> <u>units</u>

Claim: $E[(X-E[X])^2] = E\{X^2\} - (E\{X\})^2$

example: Case 2 $\quad E(X) = 5' = 60''$

$$E[X^2] = \frac{(48'')^2 + (60'')^2 + (72'')^2}{3}$$

Claim: $3696 - 60^2 = 3696 - 3600 = \underline{\underline{96}}$

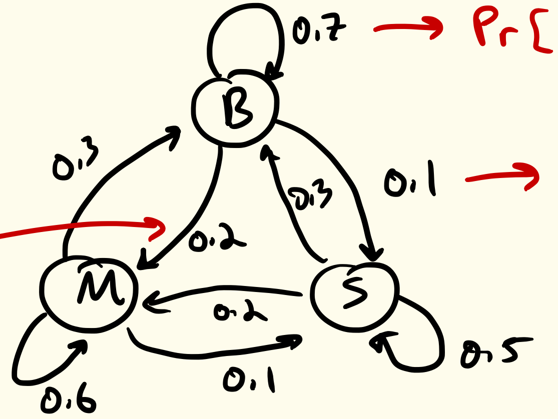
Pf:
$$\begin{aligned}
E[(X - E[X])^2] &= E[X^2 - 2X\,E[X] + (E[X])^2] \\
&= E\{X^2\} - E\{2X\,E[X]\} + E\{(E[X])^2\} \\
&= E\{X^2\} - 2E[X]\cdot E[X] + (E[X])^2 \\
&= E\{X^2\} - 2(E[X])^2 + (E[X])^2 \\
&= E\{X^2\} - (E[X])^2
\end{aligned}$$

Markov Chains

B: Bertucci's

M: Margaritas

S: Sato



$\to Pr\{B|B\}$

$Pr[M|B]$

$\to Pr[S|B]$

must sum to 1

$Pr\{B|B\} + Pr\{S|B\} + Pr[M|B] = 1$

State transition matrix:

$$P = \begin{array}{c} \\ B \\ M \\ S \end{array} \begin{array}{c} \begin{array}{ccc} B & M & S \end{array} \\ \begin{pmatrix} .7 & .2 & .1 \\ .3 & .6 & .1 \\ .3 & .2 & .5 \end{pmatrix} \end{array}$$

- all rows sum to 1
- stochastic matrix

Q: If I were to "run" the Markov Chain, what is the long-term fraction of time I would spend visiting each of the nodes?

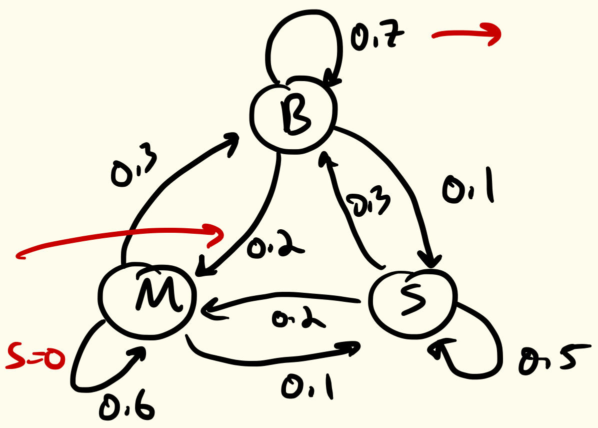
A: Stationary distribution $\vec{\pi} = \langle \pi_B, \pi_M, \pi_S \rangle \quad \pi_B + \pi_M + \pi_S = 1$

```
jay@jay-mac-2021 Downloads % ./restaurantMC.pl 10

BMBBSSSBMB

Counts: B=5; M=2; S=3

Prob: B=0.5; M=0.2; S=0.3

jay@jay-mac-2021 Downloads % ./restaurantMC.pl 100

BSBMBSSSMMMMMMMMMMSSBBBBSMMMMMSSSMMMMMBBMMMMBBBBBMMMMMMBSSSSMBBBSSMBBBBMMMMBBBMBBBMMMBSMMMMMBBBBBBBB

Counts: B=37; M=46; S=17

Prob: B=0.37; M=0.46; S=0.17

jay@jay-mac-2021 Downloads % ./restaurantMC.pl 1000

BBSBMMBMSSSSMMBBSBBSBBBMSMMMBBBBBBSSSBBBBBBBBSMMMMMMMMMMBBSBMBBBBSSBMSMMBBMMMMBSSMMMMBBBBBBBBBSSBBBMMBBMBMMBMMSSB
BSSSSBBBBMBBMMMMMSSSSSBSSSSBBMMMMBBMMBBBBBBMBBBSBBBSBBBBMMBMMBBMBMMBMMMMBBMMMMMBBBBBMBSSBBBSSMSBBBBBBBMMMMMMBBBMB
BBBMMMMBBBBBBMMMMMMMMBBBSSSMMMMSBBBMBMMMMBMBMMBSSSSSSMMSBMMBBMMMBBBBBMMMMMMBBBBBMBBBBMMMBBBBBBBMMMMMMMMMBBBBBMMBB
MMMSBMMSSMSSBBBBMSSBSBBBBBSMMBBMSBBMMBMBMMMBMMMMMMBBBBBBBBMMMBBBSSSSBBBBMMMBBBSSMMMMBBBBBBBBBBMMBMMBMMMMMBBBBBBBM
BBSSMMMMBSSBSSBMMMMMMMBBMMBBBBMSMBSSSMMBBMBBBBBSSMMBBBBBBBBBBBMMMMMMMBBMMBBMMBSSMMMMMMMBMSSSBBSMMMBBBBBBBBBMMBSSS
SBBBBBBBMSSMMMMMSSSSBBBMMBBBSSSMMMMMMBBBBBBBBBBBBBBBBBBBBMMMMMMMMMBBBBMMMMMMMMMMSBBBBSMMMMMSSMMBMMBBBBSBBBBBMSMM
BBBBBBBSMMBBBBBBBBBBBBSBBBSBSBBBBBMMMMMBMMMBBBBBBBBBBBBBBBBBBBBBMMMBBSMSBMBBBBBBMMBBBBBBBBBMBBBBBBBBBBBBBMSMMMMMMMMM
MMBMBSSBBBBMMMBBBBMBBSSBBBBBMMSSSMMMMMMMSSBBMMMBBMSMMMBMMBBBBBBSBBBBMMMBMMMMBBBBBBBBBBBBBBBBBBBBBBMBMBMMMMSMMMMBMB
BBBBBBBSSSMMSSBMMBBBBMMMSSSSMMMMSSSSMBMMMBBBSMMBBBBBBBBBBBSSBMMMMBBMMMMMMMBBBMBMMMMMBBBBBBSSBSSSM

Counts: B=488; M=366; S=146

Prob: B=0.488; M=0.366; S=0.146
```

⇒ way too long to converge, and just for 3 nodes!

① simulate — takes too long to converge

② solve algebraically



① $B = .7 \cdot B + .3 \cdot M + .3 \cdot S$

② $M = 0.2 \cdot B + 0.6\, M + 0.2 \cdot S \rightarrow -.2B + .4M - 0.2S = 0$

③ $S = 0.1 \cdot B + 0.1\, M + 0.5 \cdot S \rightarrow -.1B - .1M + .5S = 0$

① $.3B - .3M - .3S = 0$

②+③ $.3B - .3M - .3S = 0$

But also have

④ $B + M + S = 1$

$B + M + S = 1$

$5 \times ②\quad -B + 2M - S = 0 \rightarrow 3M = 1 \Rightarrow M = 1/3$

$10 \times ③\quad -B - M + 5S = 0 \rightarrow 6S = 1 \Rightarrow S = 1/6$

$B = 1/2$

Works, but in general, solving n equations in n unknowns takes $O(n^3)$ time, which is infeasible for large n.

Guess an answer for stationary distribution

$B_0 = 1/3$

$M_0 = 1/3$

$S_0 = 1/3$

① $B = .7 \cdot B + .3 \cdot M + .3 \cdot S$

② $M = 0.2 \cdot B + 0.6\, M + 0.2 \cdot S$

③ $S = 0.1 \cdot B + 0.1\, M + 0.5 \cdot S$

Try it

new guess

$B_1 = .7 \times 1/3 + .3 \times 1/3 + .3 \times 1/3 = .433$

$M_1 = .2 \times 1/3 + .6 \times 1/3 + .2 \times 1/3 = .333$

$S_1 = .1 \times 1/3 + .1 \times 1/3 + .5 \times 1/3 = .233$

① 1/2

② 1/3

③ 1/6

$B_2 = .7 \times .43\bar{3} + .3 \times .3\bar{3} + .3 \times .2\bar{3} = .473\bar{3}$

$M_2 = \ldots = .33\bar{3}$

$S_2 = \ldots \_\_\_ = .193\bar{3}$

new guess

$\Rightarrow$ iterate until convergence

```
jay@jay-mac-2021 Downloads % ./markovIterate.pl 20 transition.txt
Processing transition matrix...

0.333333333333333 0.333333333333333 0.333333333333333
0.433333333333333 0.333333333333333 0.233333333333333
0.473333333333333 0.333333333333333 0.193333333333333
0.489333333333333 0.333333333333333 0.177333333333333
0.495733333333333 0.333333333333333 0.170933333333333
0.498293333333333 0.333333333333333 0.168373333333333
0.499317333333333 0.333333333333333 0.167349333333333
0.499726933333333 0.333333333333333 0.166939733333333
0.499890773333333 0.333333333333333 0.166775893333333
0.499956309333333 0.333333333333333 0.166710357333333
0.499982523733333 0.333333333333333 0.166684142933333
0.499993009493333 0.333333333333333 0.166673657173333
0.499997203797333 0.333333333333333 0.166669462869333
0.499998881518933 0.333333333333333 0.166667785147733
0.499999552607573 0.333333333333333 0.166667114059093
0.499999821043029 0.333333333333333 0.166666845623637
0.499999928417212 0.333333333333333 0.166666738249455
0.499999971366884 0.333333333333333 0.166666695299782
0.499999988546754 0.333333333333333 0.166666678119913
0.499999995418701 0.333333333333333 0.166666671247965
0.49999999816748 0.33333333333333 0.16666666849918
```

$\Rightarrow$ rapid convergence; very efficient