CSE 312

Foundations of Computing II

Lecture 6: Chain Rule and Independence

PAUL G. ALLEN SCHOOL
OF COMPUTER SCIENCE & ENGINEERING

Anna R. Karlin

Slide Credit: Based on Stefano Tessaro's slides for 312 19au
incorporating ideas from Alex Tsun, Rachel Lin, Hunter Schafer & myself ☺

Last Class:

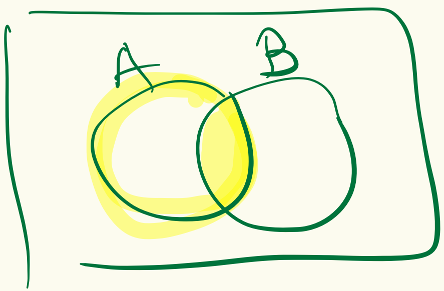
- Conditional Probability
- Bayes Theorem
- Law of Total probability

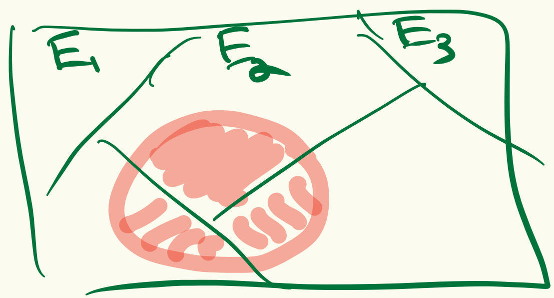
$$\boxed{\mathbb{P}(\mathcal{B}|\mathcal{A}) = \frac{\mathbb{P}(\mathcal{A} \cap \mathcal{B})}{\mathbb{P}(\mathcal{A})}}$$

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\mathbb{P}(B)}$$

$$\mathbb{P}\,(F) = \sum_{i=1}^{n} \mathbb{P}(F|E_i)\mathbb{P}(E_i)$$ E_i partition Ω

$$= \sum_{i=1}^{n} Pr(E_i \cap F)$$





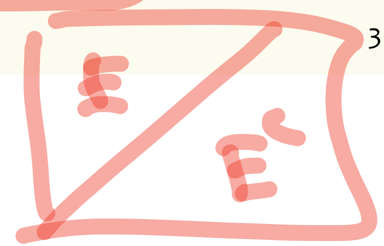
Bayes Theorem with Law of Total Probability

Bayes Theorem with LTP: Let $E_1, E_2, \ldots, E_n$ be a partition of the sample space, and F and event. Then,

$$P(E_1|F) = \frac{P(F|E_1)P(E_1)}{P(F)} = \frac{P(F|E_1)P(E_1)}{\sum_{i=1}^{n} P(F|E_i)P(E_i)}$$

Simple Partition: In particular, if E is an event with non-zero probability, then

$$P(E|F) = \frac{P(F|E)P(E)}{P(F|E)P(E) + P(F|E^C)P(E^C)}$$



Example – Zika Testing

Usually no or mild symptoms (rash); sometimes severe symptoms (paralysis).

During pregnancy: may cause birth defects.

Suppose you took a Zika test, and it returns "positive", what is the likelihood that you actually have the disease?

- Tests for diseases are rarely 100% accurate.

This example and following slides are from Lisa Yan (Stanford).

Z: have Zika
T: tests positive.

Example – Zika Testing

Suppose we know the following Zika stats

- A test is 98% effective at detecting Zika ("true positive")
- However, the test yields a "false positive" 1% of the time
- 0.5% of the US population has Zika.

$Pr(T|Z) = 0.98$
$Pr(T|\bar{Z}) = 0.01$
$Pr(Z) = 0.005$

What is the probability you have Zika (event *Z*) if you test positive (event *T*).

$P(Z|T)$

https://pollev.com/ annakarlin185

A) Less than 0.25
B) Between 0.25 and 0.5
C) Between 0.5 and 0.75
D) Between 0.75 and 1

Z : have Zika
T : tests positive.

Example – Zika Testing

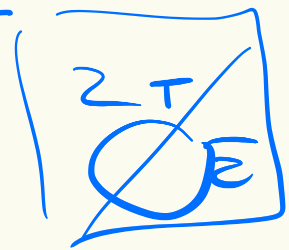
Suppose we know the following Zika stats

- A test is 98% effective at detecting Zika ("true positive")
- However, the test may yield a "false positive" 1% of the time
- 0.5% of the US population has Zika.

$\Pr(T|Z) = 0.98$
$\Pr(T|\bar{Z}) = 0.01$
$\Pr(Z) = 0.005$

What is the probability you have Zika (event *Z*) if you test positive (event *T*).

$$\Pr(Z|T) = \frac{\Pr(T|Z)\Pr(Z)}{\Pr(T)} = \frac{0.98 \cdot 0.005}{0.01485}$$



$$\Pr(T) = \underbrace{\Pr(T|Z)}_{0.98}\underbrace{\Pr(Z)}_{0.005} + \underbrace{\Pr(T|\bar{Z})}_{0.01}\underbrace{\Pr(\bar{Z})}_{\substack{1-\Pr(Z) \\ 0.995}}$$

$= 0.01485$

≈ 0.33

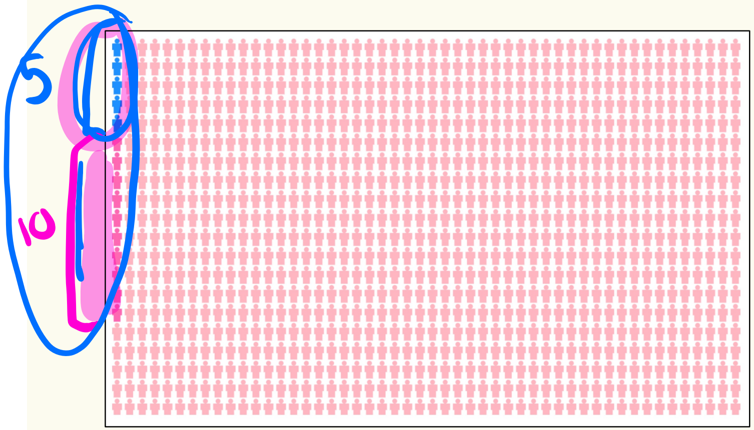
Example – Zika Testing

Have zika blue, don't pink

Suppose we know the following Zika stats

- A test is 98% effective at detecting Zika ("true positive") 100% $\Pr(T|Z) = 1$
- However, the test may yield a "false positive" 1% of the time   10/995 = approximately 1%
- 0.5% of the US population has Zika. 5% have it.

What is the probability you have Zika (event Z) if you test positive (event T).



Suppose we had 1000 people:

- 5 **have Zika** and **test positive**
- 985 **do not have Zika** and **test negative**
- 10 **do not have Zika** and **test positive**

$$\frac{5}{5+10} = \frac{1}{3} \approx 0.33$$

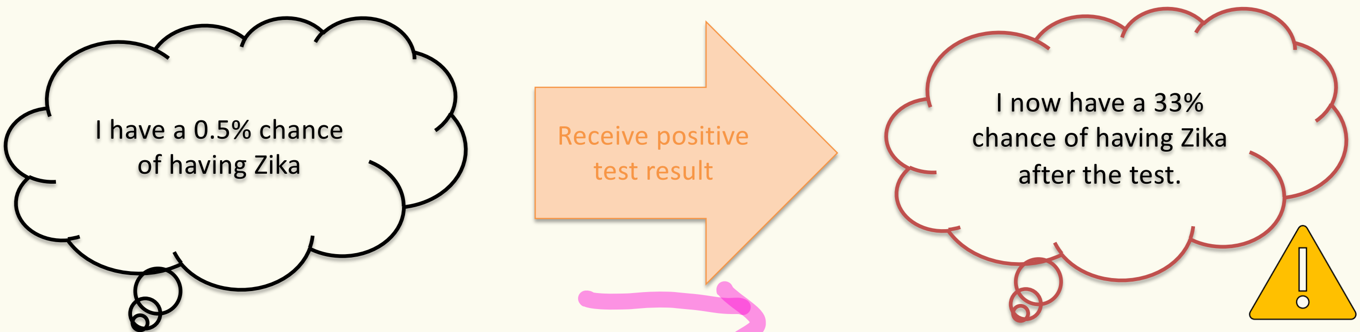
Demo

Philosophy – Updating Beliefs

While it's not 98% that you have the disease, your beliefs changed **drastically**

Z = you have Zika

T = you test positive for Zika



Prior: P(Z)

Posterior: P(Z|T)

Example – Zika Testing

Suppose we know the following Zika stats

- A test is 98% effective at detecting Zika ("true positive")
- However, the test may yield a "false positive" 1% of the time
- 0.5% of the US population has Zika.

$Pr(T|Z) = 0.98$

What is the probability you test negative (event $\bar{T}$) if you have Zika (event Z)?

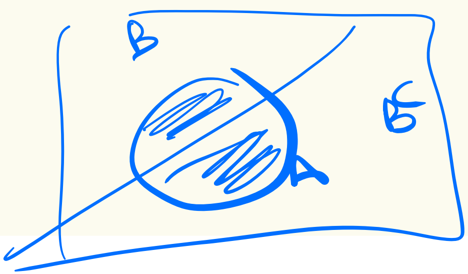
$Pr(\bar{T}|Z) = 1 - Pr(T|Z) = 0.02$

Conditional Probability Define a Probability Space

The probability conditioned on A follows the same properties as (unconditional) probability.

Example. $\mathbb{P}(\mathcal{B}^c|\mathcal{A}) = 1 - \mathbb{P}(\mathcal{B}|\mathcal{A})$

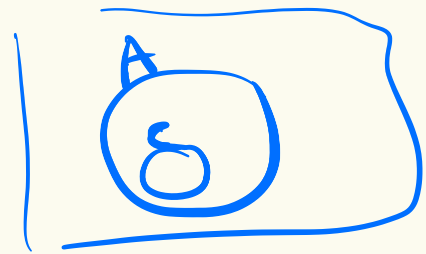
$$P(B|A) + P(B^c|A) = \frac{P(B\cap A)}{P(A)} + \frac{Pr(B^c\cap A)}{P(A)}$$



$$= \frac{P(B\cap A) + Pr(B^c\cap A)}{P(A)} = \frac{Pr(A)}{P(A)} = 1$$

Conditional Probability Define a Probability Space

The probability conditioned on A follows the same properties as Ω (unconditional) probability.



Example. $\mathbb{P}(\mathcal{B}^c|\mathcal{A}) = 1 - \mathbb{P}(\mathcal{B}|\mathcal{A})$

Formally. $(\Omega, \mathbb{P})$ is a probability space + $\mathbb{P}(\mathcal{A}) > 0$

$\Rightarrow$ $\boxed{(\mathcal{A}, \mathbb{P}(\cdot|\mathcal{A})) \text{ is a probability space}}$

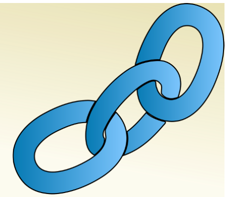
$f(\cdot)$

$\cancel{P(A|B) + P(A|\bar{B})} = ?$

Today:

- Chain Rule
- Independence
- Sequential Process

Chain Rule

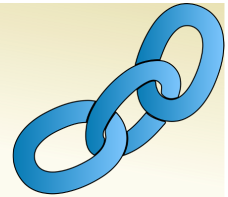


$$\mathbb{P}(\mathcal{B}|\mathcal{A}) = \frac{\mathbb{P}(\mathcal{A} \cap \mathcal{B})}{\mathbb{P}(\mathcal{A})} \Rightarrow \mathbb{P}(\mathcal{A})\mathbb{P}(\mathcal{B}|\mathcal{A}) = \mathbb{P}(\mathcal{A} \cap \mathcal{B})$$

$$P(\underbrace{A_1 \cap A_2}_{A} \cap \underbrace{A_3}_{B}) = Pr(A_3 | A_1 \cap A_2)\underbrace{Pr(A_1 \cap A_2)}$$

$$= Pr(A_3 | A_1 \cap A_2)\, Pr(A_2 | A_1)\, Pr(A_1)$$

Chain Rule



$$\mathbb{P}(\mathcal{B}|\mathcal{A}) = \frac{\mathbb{P}(\mathcal{A} \cap \mathcal{B})}{\mathbb{P}(\mathcal{A})} \Rightarrow \mathbb{P}(\mathcal{A})\mathbb{P}(\mathcal{B}|\mathcal{A}) = \mathbb{P}(\mathcal{A} \cap \mathcal{B})$$

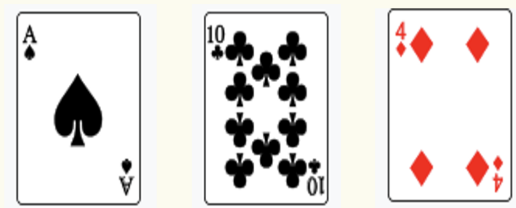
Theorem. (Chain Rule) For events $\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_n$,

$$\mathbb{P}(\mathcal{A}_1 \cap \cdots \cap \mathcal{A}_n) = \mathbb{P}(\mathcal{A}_1) \cdot \mathbb{P}(\mathcal{A}_2|\mathcal{A}_1) \cdot \mathbb{P}(\mathcal{A}_3|\mathcal{A}_1 \cap \mathcal{A}_2) \cdots \mathbb{P}(\mathcal{A}_n|\mathcal{A}_1 \cap \mathcal{A}_2 \cap \cdots \cap \mathcal{A}_{n-1})$$

An easy way to remember: We have n tasks and we can do them sequentially, conditioning on the outcome of previous tasks

Chain Rule Example

Have a Standard 52-Card Deck. Shuffle It, and draw the top 3 cards **in order**. (uniform probability space).

What is P () = P(**A** ∩ **B** ∩ **C**)?

A: Ace of Spades First
B: 10 of Clubs Second
C: 4 of Diamonds Third

Pr(AS first) Pr(10 Clubs 2nd | AS first) Pr(4♢ | AS first ∩ 10 Clubs second)

$\frac{1}{52}$ $\frac{1}{51}$ $\frac{1}{50}$

$= \frac{\#\text{perm AS first}}{\#} = \frac{51!}{52!}$

$Pr(10C \text{ 2nd} \mid AS \text{ first}) = \frac{Pr(10C \text{ 2nd} \cap AS \text{ first})}{Pr(AS \text{ first})}$

50!

$= \frac{50!}{51!} = \frac{1}{51}$

AS

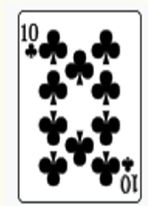
Chain Rule Example

$Pr(10 \text{ Clubs } 2nd) = \frac{1}{52}$

10C

Have a Standard 52-Card Deck. Shuffle It, and draw the top 3 cards in order. (uniform probability space).

What is P (



) = P(**A** ∩ **B** ∩ **C**)?

A: Ace of Spades First

B: 10 of Clubs Second

C: 4 of Diamonds Third

$$\mathbb{P}(A) \cdot \mathbb{P}(B|A) \cdot \mathbb{P}(C|A \cap B)$$

$$\frac{1}{52} \cdot \frac{1}{51} \cdot \frac{1}{50}$$

Independence

Definition. Two events $\mathcal{A}$ and $\mathcal{B}$ are (statistically) **independent** if

$$\mathbb{P}(\mathcal{A} \cap \mathcal{B}) = \mathbb{P}(\mathcal{A}) \cdot \mathbb{P}(\mathcal{B}).$$

Alternatively,

- If $\mathbb{P}(\mathcal{A}) \neq 0$, equivalent to $\boxed{\mathbb{P}(\mathcal{B}|\mathcal{A}) = \mathbb{P}(B)}$
- If $\mathbb{P}(\mathcal{B}) \neq 0$, equivalent to $\mathbb{P}(\mathcal{A}|\mathcal{B}) = \mathbb{P}(\mathcal{A})$

$$Pr(B|A) \triangleq \frac{P(A \cap B)}{P(A)} = P(B)$$

"The probability that $\mathcal{B}$ occurs after observing $\mathcal{A}$" -- Posterior
= "The probability that $\mathcal{B}$ occurs" -- Prior

Example -- Independence

Toss a coin 3 times. Each of 8 outcomes equally likely.

- A = {at most one T} = {HHH, $\boxed{\text{HHT, HTH, THH}}$}
- B = {at most 2 Heads}= {HHH}^{c}

Independent?

$$\mathbb{P}(\mathcal{A} \cap \mathcal{B}) \stackrel{?}{=} \mathbb{P}(\mathcal{A}) \cdot \mathbb{P}(\mathcal{B})$$

X

$Pr(A) = \frac{1}{2}$

$Pr(B) = \frac{7}{8}$

$Pr(A \cap B) = \frac{3}{8}$

Poll:

A. Yes, independent

B. No

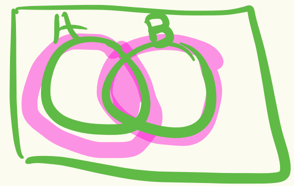
pollev/annakarlin185

Often probability space $(\Omega, \mathbb{P})$ is **defined** using independence

Example – Network Communication

$\Pr(E \cup F) = P(E) + P(F) - P(E \cap F)$

Each link works with the probability given, **independently**.
What's the probability A and D can communicate?

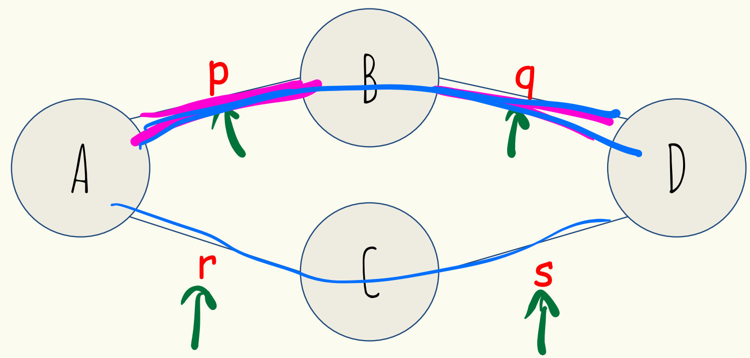


$\mathbb{P}(AD) = ?$

$= \Pr(\overset{E}{AB \cap BD} \cup \overset{F}{AC \cap CD})$

top bottom

$= \Pr(\underbrace{AB \cap BD}_{p \cdot q}) + P(\underbrace{AC \cap CD}_{r \cdot s}) - \Pr(\underbrace{AB \cap BD \cap AC \cap CD}_{p \cdot q \cdot r \cdot s})$



— mutually excl events
disjoint events

Can mutually exclusive events be i..

A B

$P(A \cap B) = 0$

— indep events A, B

$P(A \cap B) = P(A) \cdot P(B)$

Example – Network Communication

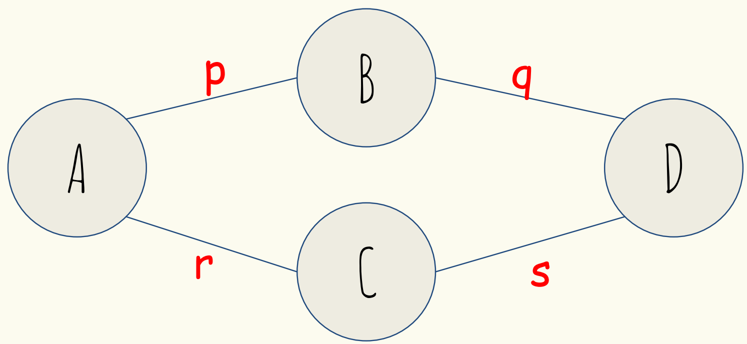
Each link works with the probability given, **independently**.
What's the probability A and D can communicate?

$$\mathbb{P}(AD) = \mathbb{P}(AB \cap BD \text{ } or \text{ } AC \cap CD)$$

$$= \mathbb{P}(AB \cap BD) + \mathbb{P}(AC \cap CD) - \mathbb{P}(AB \cap BD \cap AC \cap CD)$$

$$\mathbb{P}(AB \cap BD) = \mathbb{P}(AB) \cdot \mathbb{P}(BD) = pq$$

$$\mathbb{P}(AC \cap CD) = \mathbb{P}(AC) \cdot \mathbb{P}(CD) = rs$$



$$\mathbb{P}(AB \cap BD \cap AC \cap CD) = \mathbb{P}(AB) \cdot \mathbb{P}(BD) \cdot \mathbb{P}(AC) \cdot \mathbb{P}(CD) = pqrs$$

Example – Biased coin

We have a biased coin comes up Heads with probability 2/3; Each flip is independent of all other fips. Suppose it is tossed 3 times.

$$\mathbb{P}(HHH) = Pr(H)Pr(H)Pr(H) = \left(\frac{2}{3}\right)^3$$

$$\mathbb{P}(TTT) = Pr(T)Pr(T)P(T) = \left(\frac{1}{3}\right)^3$$

$$\mathbb{P}(HTT) = Pr(H)P(T)P(T) = \frac{2}{3}\left(\frac{1}{3}\right)^2$$

$Pr(A \cap B) = 0$

$\boxed{Pr(B) = 0}$

Example – Biased coin

We have a biased coin comes up Heads with probability 2/3, independently of other flips. Suppose it is tossed 3 times.

$\mathbb{P}(2\ heads\ in\ 3\ tosses) =$

$= P(HHT, HTH, THH)$

https://pollev.com/ annakarlin185

A) $(2/3)^2\ 1/3$
B) $2/3$
C) $\boxed{3\ (2/3)^2\ 1/3}$ ✓
D) $(1/3)^2$