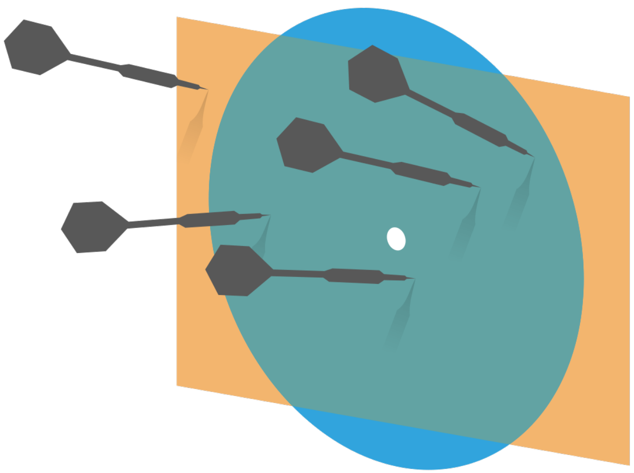
Probability and Statistics for Computer Science



Credit: wikipedia

"A major use of probability in statistical inference is the updating of probabilities when certain events are observed" – Prof. M.H. DeGroot

Hongye Liu, Teaching Assistant Prof, CS361, UIUC, 9.8.2021

Warm up

1) Ways of forming a queue with 10 students

$10 \times 9 \times 8 \times \cdots 1 = 10!$

2) Ways of forming a queue with 5 students out of 10 students

$\frac{10!}{5!}$ $\frac{N!}{(N-K)!}$

$10 \times 9 \times 8 \times 7 \times 6 \times \frac{5!}{5!}$

3) Ways of forming a team of 5 from 10 students.

$\frac{10!}{5!\,5!} = \frac{N!}{K!(N-K)!}$

Which is larger?

1) $\binom{120}{90}$ 2) $\binom{120}{30}$

$\frac{120!}{90!\,30!}$ $\frac{120!}{30!\,90!}$

A. 1)

B. 2)

$\boxed{\text{C.}}$ None

$\binom{N}{K} = \binom{N}{N-K}$

Objectives

- Probability
  - More probability calculation
  - Conditional Probability
    - Bayes rule
    - Independence

Senate Committee problem

The United States Senate contains **two** senators from each of the **50** states. If a committee of **eight** senators is selected at random, what is the probability that it will contain at least one of the two senators from **IL**?

E^c: None of IL senators got in

$P(E) = 1 - P(E^c)$

$|E^c| = \binom{98}{8}\binom{2}{0}$

$|\Omega| = \binom{100}{8}$

$$P(E) = P(E_a) + P(E_b)$$

1 sen. from IL

2 Sen from IL

$$\frac{|E_a|}{|\Omega|} = \frac{\binom{98}{7} \cdot \binom{2}{1}}{\binom{100}{8}}$$

$$|E_b| = \binom{98}{6}\binom{2}{2}$$

Probability: Birthday problem

- Among 30 people, what is the probability that at least 2 of them celebrate their birthday on the same day? Assume that there is no February 29 and each day of the year is equally likely to be a birthday.

$\underbrace{365 \times 364 \times \cdots}_{30} = P\binom{365}{30}$

E^c :

$|E^c| =$

$P(E) = 1 - P(E^c)$

$|\Omega| = 365^{30}$

$P(E^c) = \frac{|E^c|}{|\Omega|} = \frac{P\binom{365}{30}}{365^{30}}$

2/2 2/2

2/2 2/1

2/1 2/2

Conditional Probability

- Motivation of conditional probability

If a person in CS tested positive for covid, how likely does this person have the disease?

T : test positive

D : has the disease

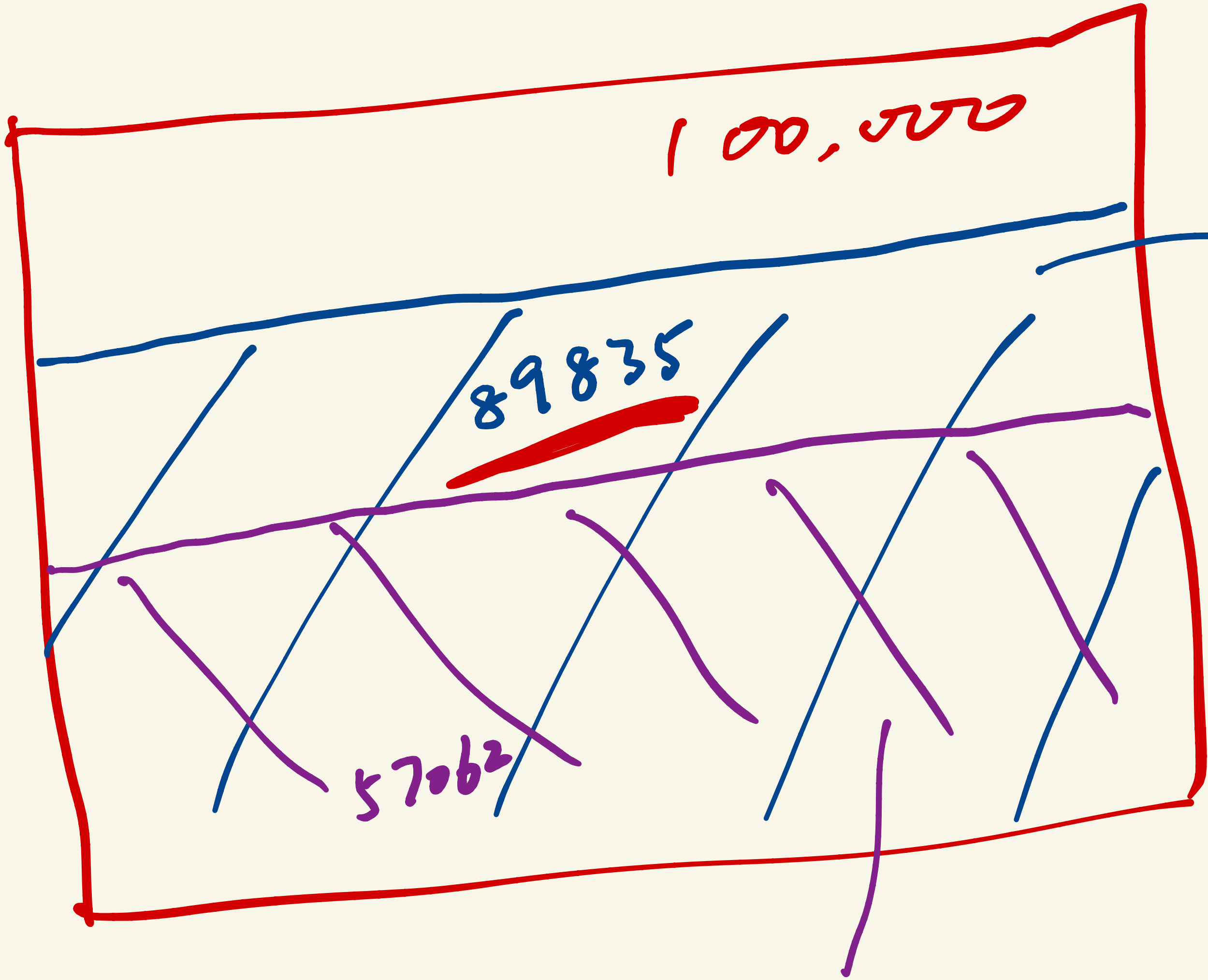
$P(D|T)$

Conditional Probability

- Example:

An insurance company knows in a population of 100 thousands females, 89.835% expect to live to age 60, while 57.062% can expect to live to 80. Given a woman at the age of 60, what is the probability that she lives to 80?

$|\Omega| = 100,000$



$|E_{60}| = 89,835$

$|E_{80}| = 57,062$

$$P(E_{80} \mid E_{60}) = \frac{|E_{80}|}{|\Omega_n|} = \frac{57,062}{89,835}$$

$|\Omega_n| = ?$ 89,835

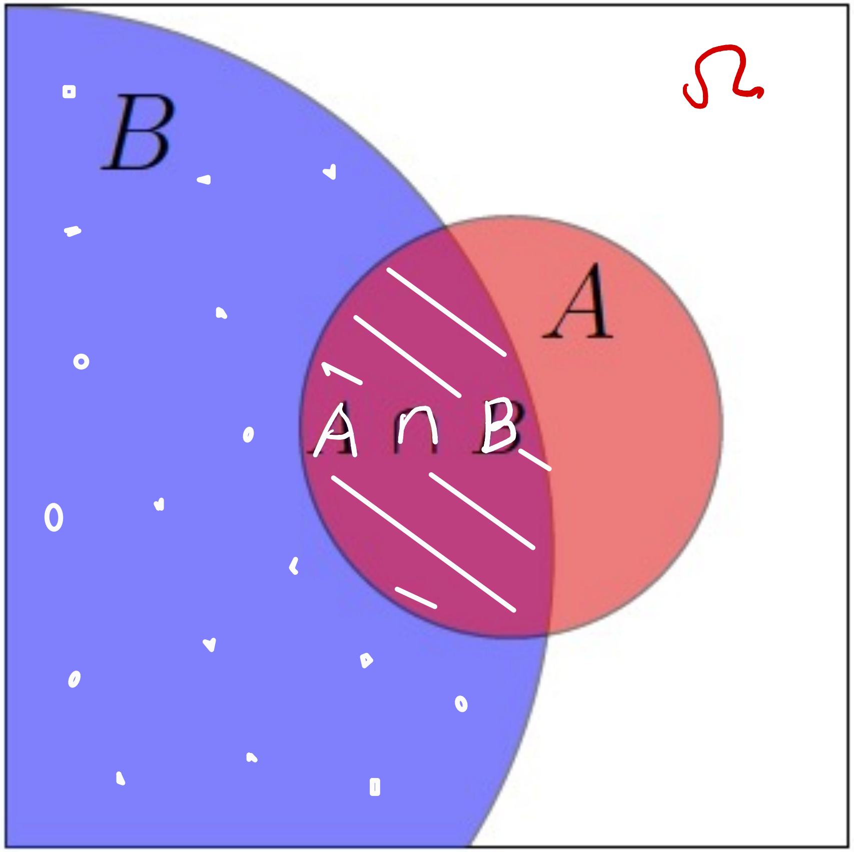
Conditional Probability

- Given the condition she is 60 already, the size of the sample space for the outcomes has changed to

  89,835 instead of 100,000

Conditional Probability

- The probability of ***A*** given ***B***



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

$$P(B) \neq 0$$

The "Size" analogy

Credit: Prof. Jeremy Orloff & Jonathan Bloom

Conditional Probability

A : a woman lives to 80

$$P(A|B) = \frac{57,062}{89,835} = 0.6352$$

B : a woman is at 60 now

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

$$= \frac{57,062 / 100K}{89,835 / 100K}$$

While $P(A) = \frac{57,062}{100,000} = 0.57062$

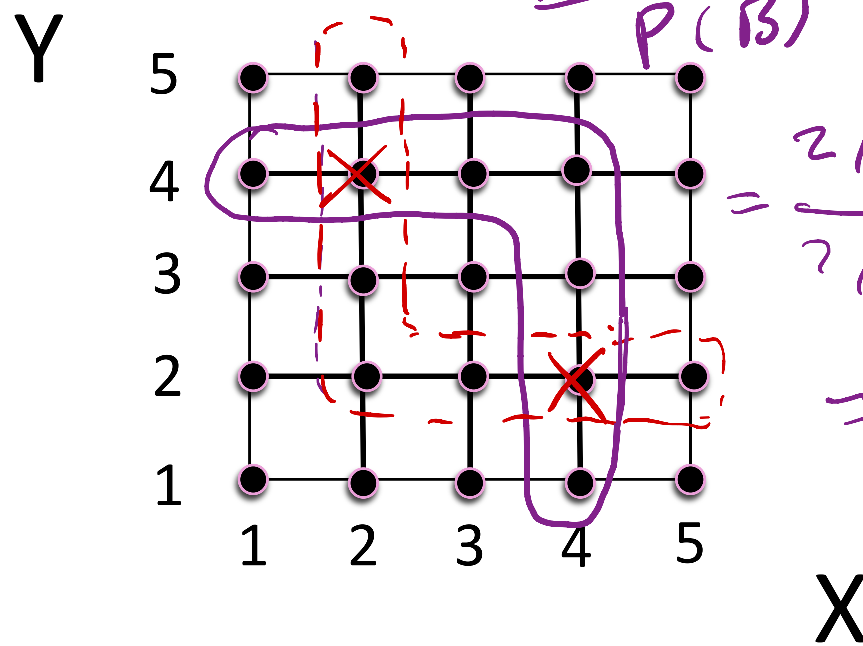
Conditional Probability: die example

Throw 5-sided fair die twice.

$$A : \underline{max(X, Y)} = 4$$

$$B : \underline{min(X, Y)} = 2$$

$$P(A|B) = ?$$



$$\frac{P(A \cap B)}{P(B)} = \frac{2/25}{7/25} = \frac{2}{7}$$

$$P(A|B) = \frac{|E| = 2}{|\Omega| = 7} = \frac{2}{7}$$

Conditional probability, that is?

$P(\Omega) = 1$

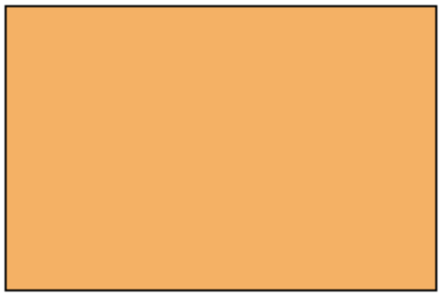
$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad P(B) \neq 0$$

Still a probability! It satisfies the three axioms

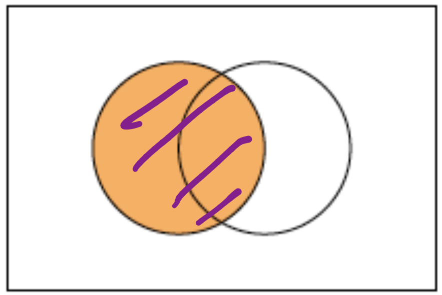
$P(A|B) + P(A^c|B) = ?\ 1$

$P(A_1 \cup A_2|B) = ?$ if $A_1 \cap A_2 = \phi$

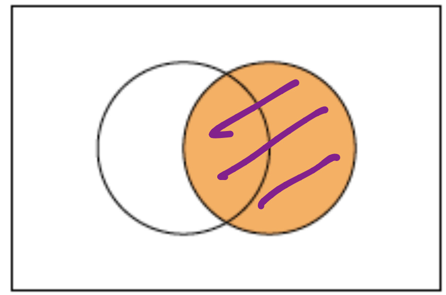
Venn Diagrams of events as sets



Ω

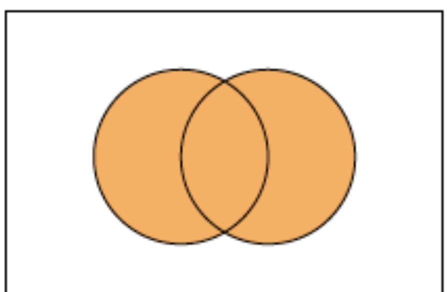


E_1

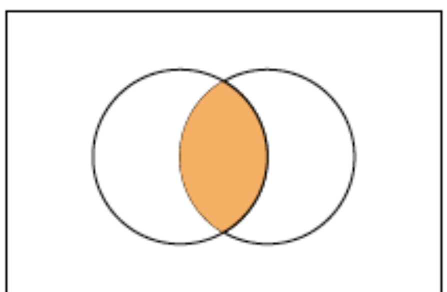


E_2

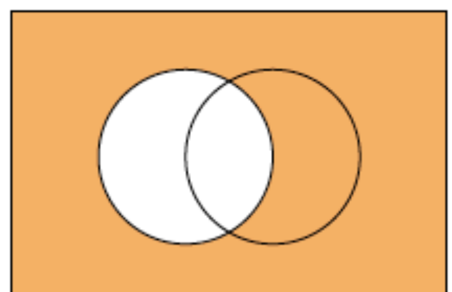
$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$



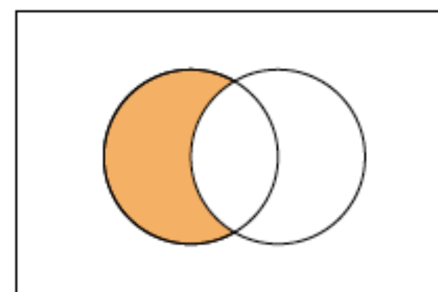
$E_1 \cup E_2$



$E_1 \cap E_2$



${E_1}^c$



$E_1 - E_2$

Multiplication rule using conditional probability

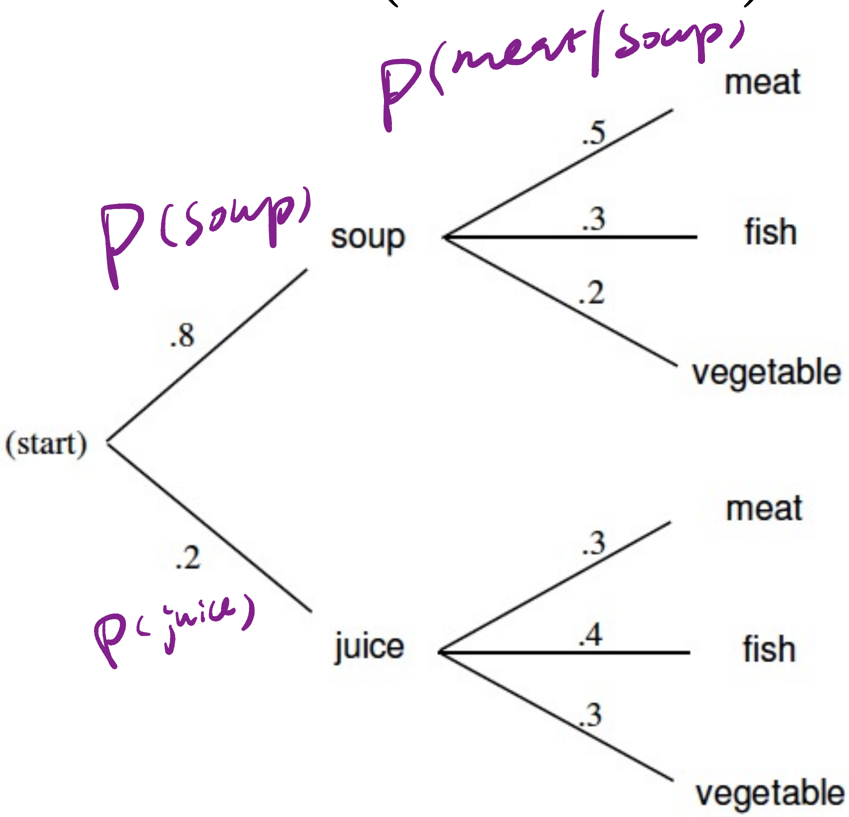
- Joint event

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

$$\Rightarrow \boxed{P(A \cap B) = P(A|B)\underbrace{P(B)}_{\text{prior}}}$$

Multiplication using conditional probability

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



$$P(soup \cap meat) =$$
$$P(meat|soup)P(soup)$$
$$= 0.5 \times 0.8 = 0.4$$

Simulation of Conditional Probability

http://www.randomservices.org/random/apps/ConditionalProbabilityExperiment.html

Additional References

- Charles M. Grinstead and J. Laurie Snell "Introduction to Probability"
- Morris H. Degroot and Mark J. Schervish "Probability and Statistics"

Assignments

- Work on Module Week 3 on Canvas
- Next time: More on independence and conditional probability

Laws of Sets

Commutative Laws

$A \cap B = B \cap A$
$A \cup B = B \cup A$

Associative Laws

$(A \cap B) \cap C = A \cap (B \cap C)$
$(A \cup B) \cup C = A \cup (B \cup C)$

Distributive Laws

$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

Laws of Sets

Idempotent Laws

$A \cap A = A$
$A \cup A = A$

Identity Laws

$A \cup \emptyset = A$
$A \cap U = A$
$A \cup U = U$
$A \cap \emptyset = \emptyset$

Involution Law $(A^c)^c = A$

U is the complete set

Complement Laws

$A \cup A^c = U$
$A \cap A^c = \emptyset$
$U^c = \emptyset$
$\emptyset^c = U$

De Morgan's Laws

$(A \cap B)^c = A^c \cup B^c$
$(A \cup B)^c = A^c \cap B^c$

See you next time

See You!