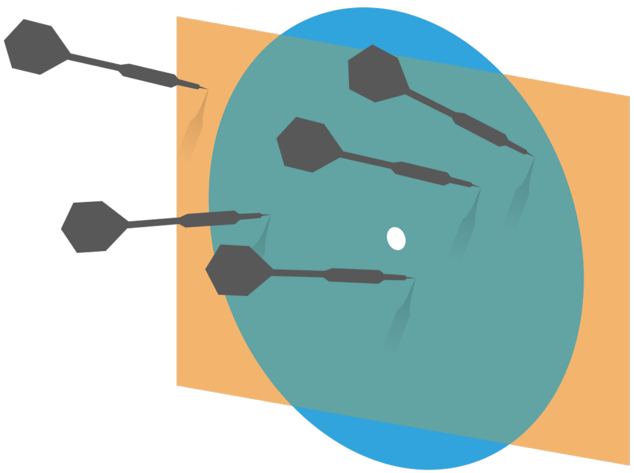
Probability and Statistics for Computer Science



Credit: wikipedia

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

Warm up

1) Ways of forming a queue with 10 students

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

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

$10!/5! = \binom{10}{5} \cdot 5!$ $\qquad 10 \times 9 \times 8 \times 7 \times 6$

3) Ways of forming committees of 5 from 10 students

$\binom{10}{5}$

Which is larger?

1) $\binom{110}{70}$ 2) $\binom{110}{40}$

A. 1)

B. 2)

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

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

Last time

- Probability a first look
  - Outcome and Sample Space
  - Event
  - Probability

    Probability axioms & Properties
  - Calculating probability

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 from IL

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

$P(E^c) = \frac{\binom{98}{8}}{\binom{100}{8}}$ $\begin{array}{l} \to |E^c| \\ \to |\Omega| \end{array}$

P(1 from IL) + P(2 from IL)

$$\frac{\binom{2}{1}\cdot\binom{98}{7}}{\binom{100}{8}} + \frac{\binom{2}{2}\binom{98}{6}}{\binom{100}{8}}$$

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.

E^c : No two share B-day

HY	You
2/1	2/2
2/2	2/1
...	

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

$$= \frac{365 P_{30} \rightarrow \frac{365!}{335!}}{365 \times 365 \times \cdots \rightarrow 365^{30}}$$

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

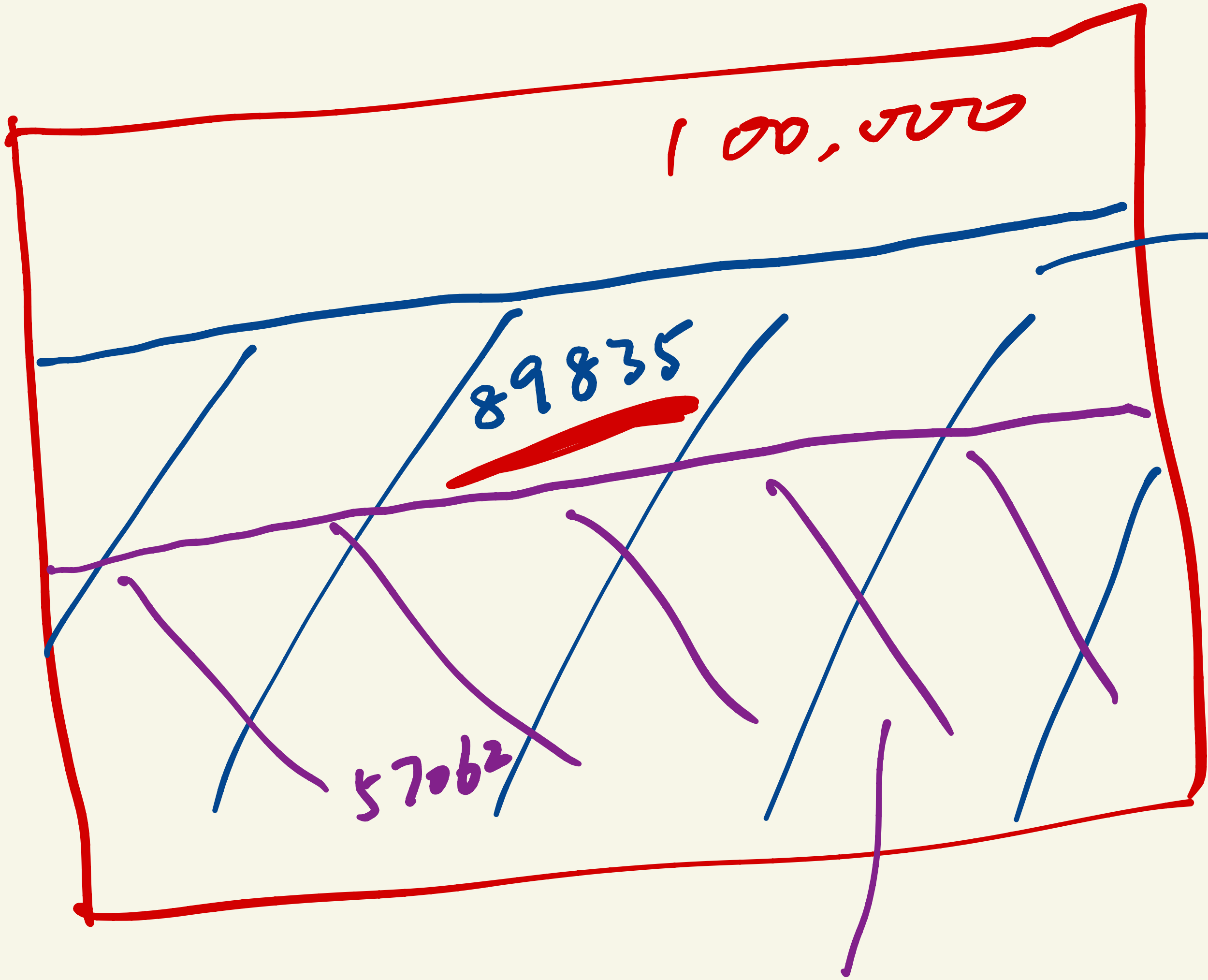
$P(D|T)$

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?

$$|\Omega| = 100,000$$



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

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

$$\underline{|\Omega_n|} = ? \quad 89,835$$

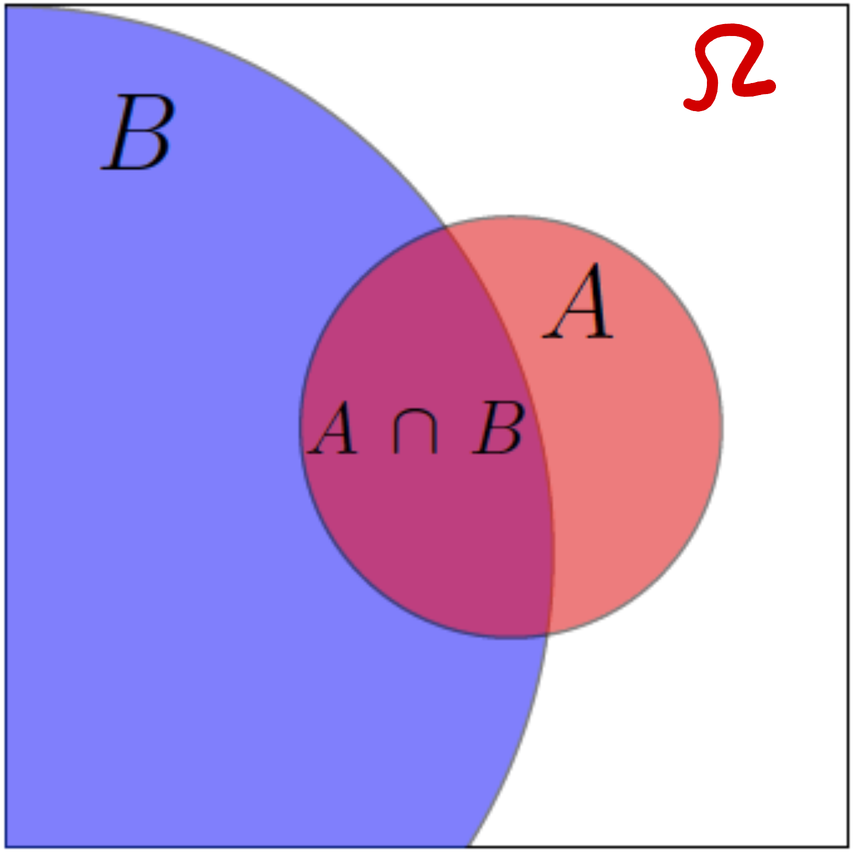
Conditional Probability

- 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 at least

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad \frac{\frac{57,062}{100,000}}{\frac{89,835}{100,000}}$$

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

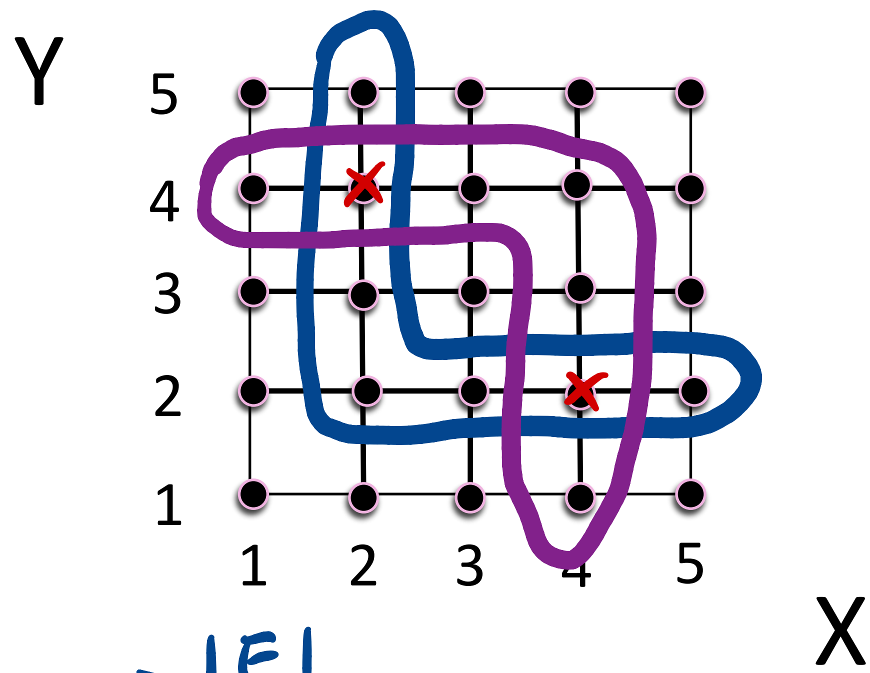
Conditional Probability: die example

Throw 5-sided fair die twice.

$$A : \underline{max(X,Y)} = 4$$
$$B : \underline{min(X,Y)} = 2$$

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

$\frac{2}{7}$ → $|E|$, → $|\Omega|$



Conditional probability, that is?

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

$$= \frac{2/25}{7/25} = \frac{2}{7}$$

Conditional probability, that is?

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

Still a probability! It satisfies the three axioms

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

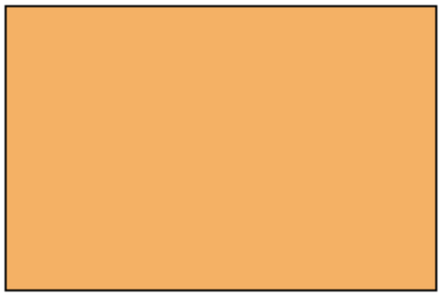
$P(A^c|B) = \frac{P(A^c \cap B)}{P(B)}$

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

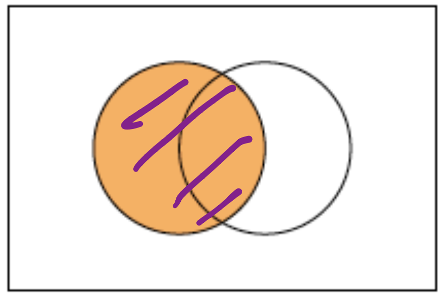
$P(A_1 \cup A_2|B) = ?$ if $A_1 \cap A_2 = \phi$

$= P(A_1|B) + P(A_2|B)$

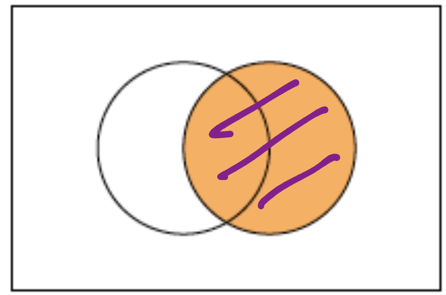
Venn Diagrams of events as sets



Ω

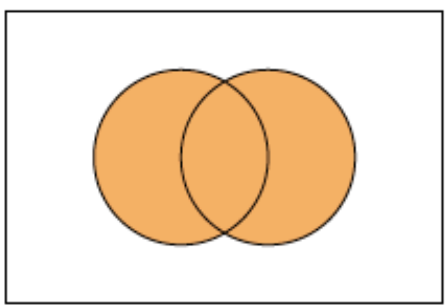


E_1

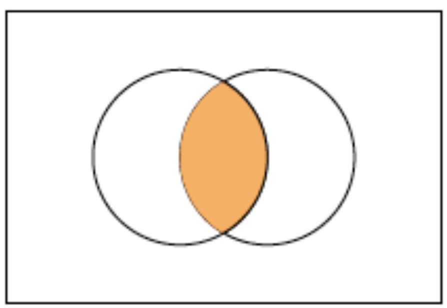


E_2

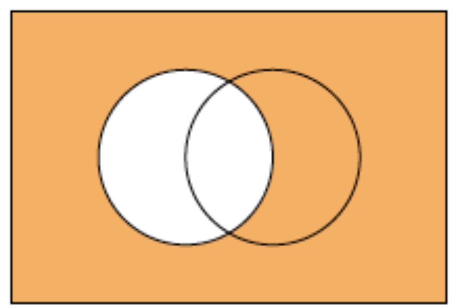
$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$



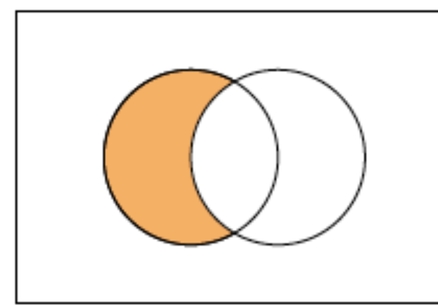
$E_1 \cup E_2$



$E_1 \cap E_2$



${E_1}^c$



$E_1 - E_2$

Multiplication rule using conditional probability

- Joint event

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

$\frac{2}{50}$ × 40%

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

Multiplication using conditional probability

P(beef | meat (soup)

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

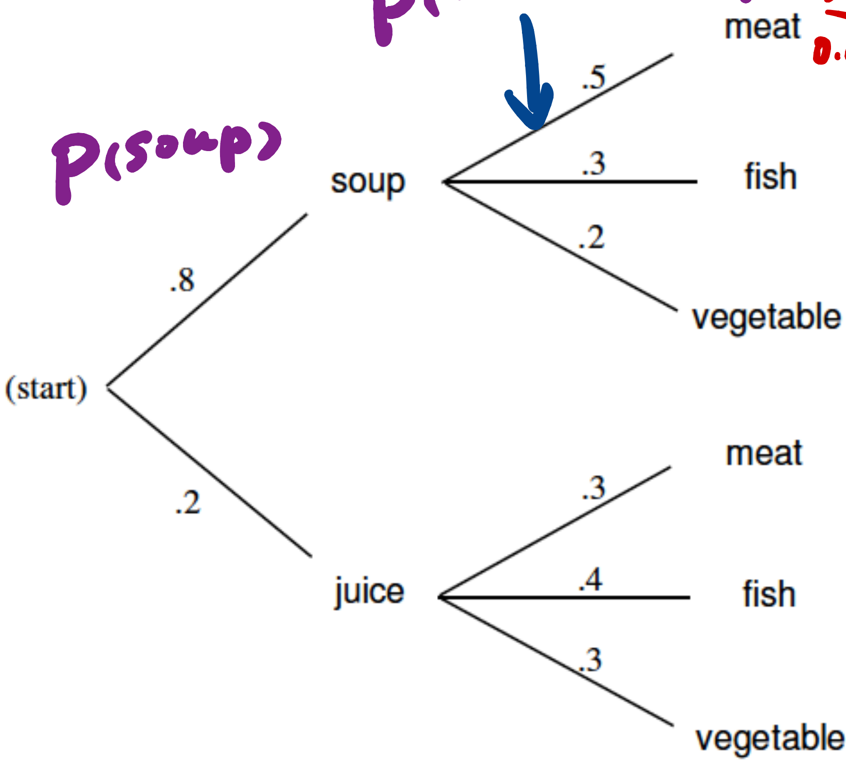
0.6 beef

0.4 pork

prior

P(meat|soup)

P(soup)



$P(A \cap B \cap C)$

soup meat beef = ?

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

$P(fish \cap juice)$

$= 0.2 \times 0.4$

Symmetry of joint event in terms of conditional prob.

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

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

Symmetry of joint event in terms of conditional prob.

$B \cap A = A \cap B$

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

$$\Rightarrow$$

$$\boxed{P(A|B)P(B) = P(B|A)P(A)}$$

The famous **Bayes** rule

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

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

Thomas Bayes (1701-1761)

Bayes rule: lemon cars

There are two car factories, **A** and **B**, that supply the same dealer. Factory **A** produced **1000** cars, of which **10** were lemons. Factory **B** produced **2** cars and both were lemons. You bought a car that turned out to be a lemon. What is the probability that it came from factory **B**?

$$P(B \mid L) = \frac{P(L \mid B) \cdot P(B)}{P(L)}$$

?

$$= \frac{1 \times \frac{2}{1002}}{12/1002} = \frac{1}{6}$$

Bayes rule: lemon cars

There are two car factories, A and B, that supply the same dealer. Factory A produced 1000 cars, of which 10 were lemons. Factory B produced 2 cars and both were lemons. You bought a car that turned out to be a lemon. What is the probability that it came from factory B?

$$P(B|L) = \frac{P(L|B)P(B)}{P(L)}$$

Bayes rule: lemon cars

Given the above information, what is the probability that it came from factory A?

$$P(A|L) = ? \qquad 1 - P(B|L)$$

Simulation of Conditional Probability

http://
www.randomservices.org/
random/apps/
ConditionalProbabilityExperim
ent.html

Commutative

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

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

Associative

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

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

Distributive

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

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

Idempotent

$A \cap A = A$

$A \cup A = A$

Identity

$A \cup \emptyset = A$

$A \cap \emptyset = \emptyset$

$A \cup \Omega = \Omega$

$A \cap \Omega = A$

Complement

$A \cup A^c = \Omega$

$A \cap A^c = \emptyset$

$\Omega^c = \emptyset$

$\emptyset^c = \Omega$

De Morgan's

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

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

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 Compass
- Next time: More on independence and conditional probability

See you next time

See You!