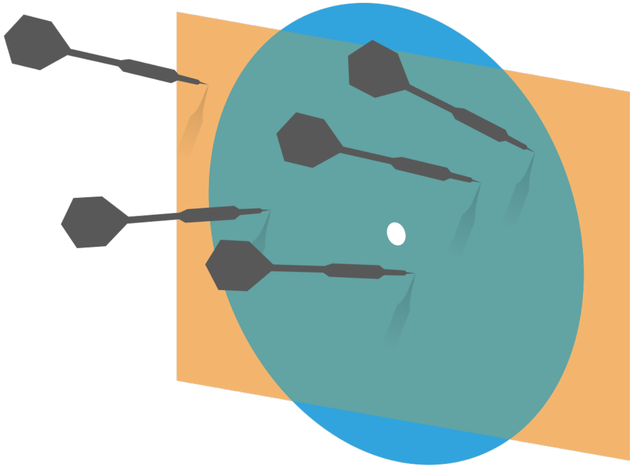
Probability and Statistics for Computer Science



Credit: wikipedia

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

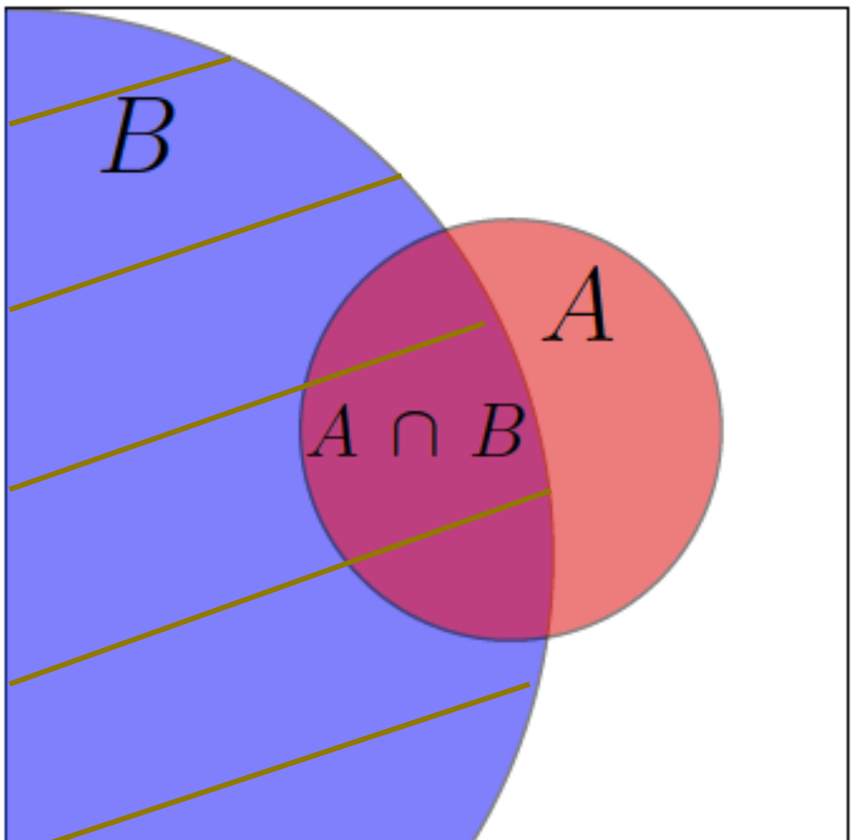
Hongye Liu, Teaching Assistant Prof, CS361, UIUC, 2.11.2021

Objectives

- Conditional Probability
  - Review
  - Total probability
  - Independence

Conditional Probability

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



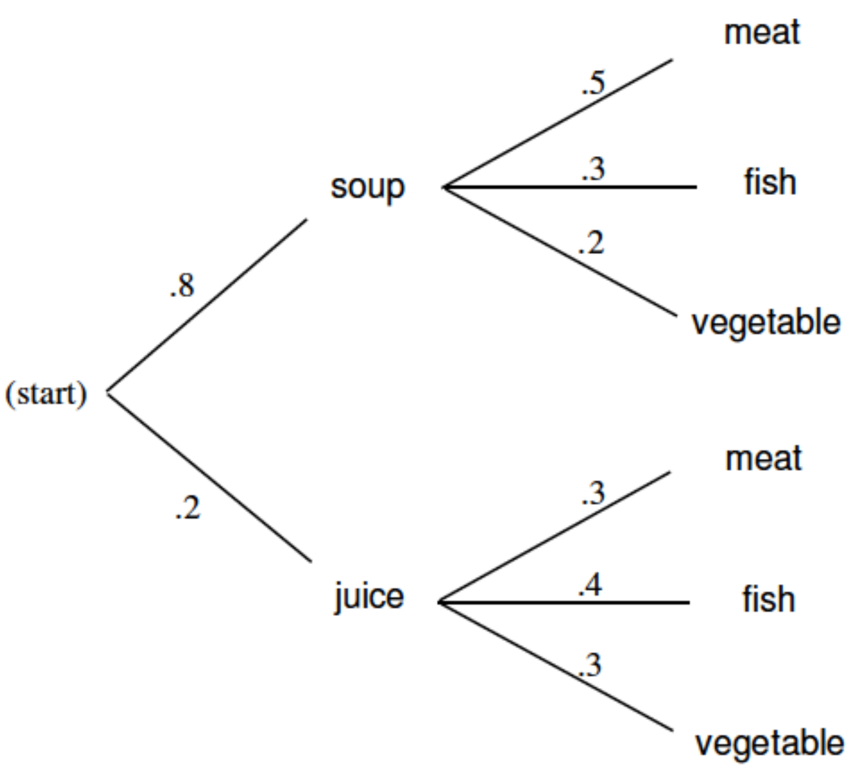
$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

$$P(B) \neq 0$$

The line-crossed area is the new sample space for conditional P(A| B)

Joint Probability Calculation

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



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

Bayes rule

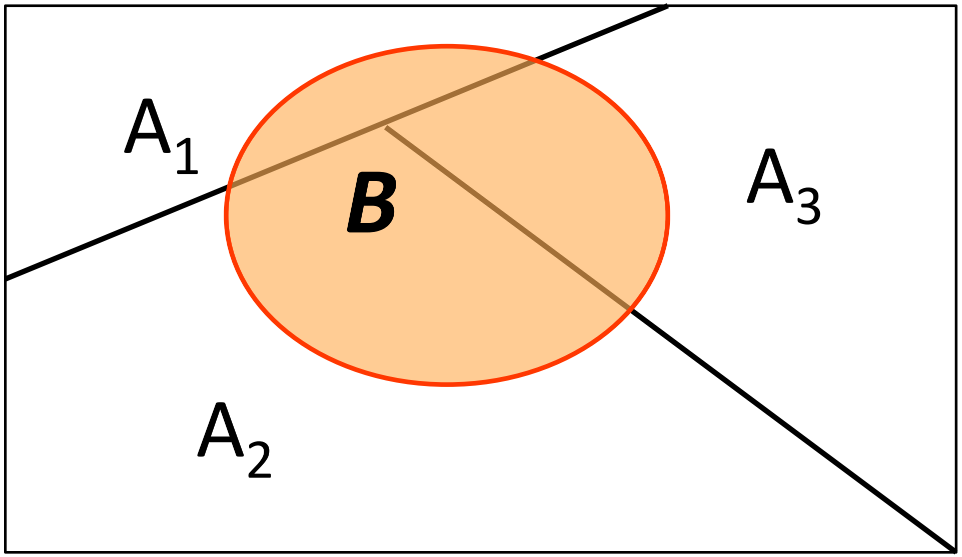
- Given the definition of conditional probability and the symmetry of joint probability, we have:

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

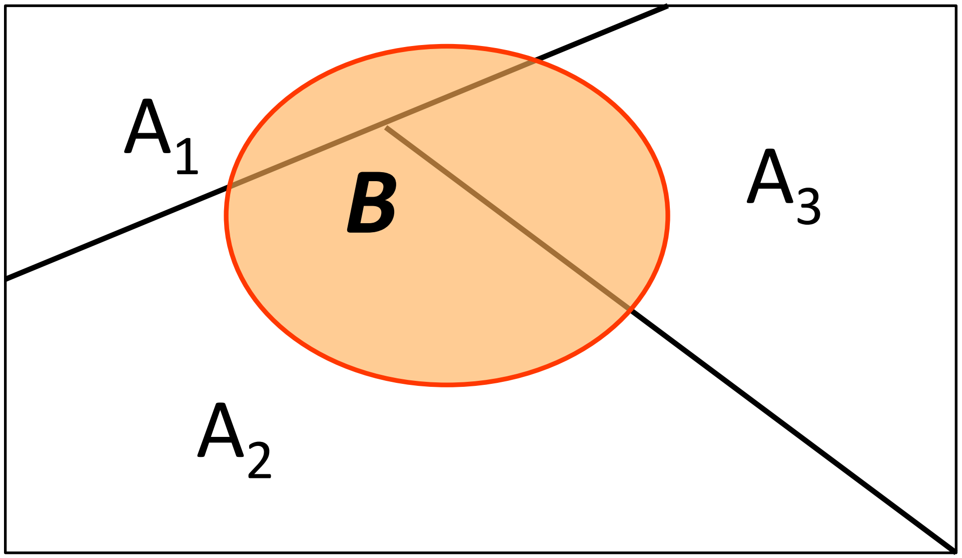
And it leads to the famous Bayes rule:

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

Total probability



Total probability general form



Total probability:

Bayes rule using total prob.

Bayes rule: rare disease test

There is a blood test for a rare disease. The frequency of the disease is 1/100,000. If one has it, the test confirms it with probability 0.95. If one doesn't have, the test gives false positive with probability 0.001. What is $P(D|T)$, the probability of having disease given a positive test result?

$$P(D|T) = \frac{P(T|D)P(D)}{P(T)}$$

Using total prob.

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

Bayes rule: rare disease test

There is a blood test for a rare disease. The frequency of the disease is **1/100,000**. If one has it, the test confirms it with probability 0.95. If one doesn't have, the test gives false positive with probability **0.001**. What is $P(D|T)$, the probability of having disease given a positive test result?

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

Independence

- One definition:

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

Whether A happened doesn't change the probability of B and vice versa

Independence: example

- Suppose that we have a fair coin and it is tossed twice. let A be the event “the first toss is a head” and B the event “the two outcomes are the same.”

- These two events are independent!

Independence

- Alternative definition

LHS by definition

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

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

$$\Rightarrow \boxed{P(A \cap B) = P(A)P(B)}$$

Testing Independence:

- Suppose you draw one card from a standard deck of cards. E_1 is the event that the card is a King, Queen or Jack. E_2 is the event the card is a Heart. Are E_1 and E_2 independent?

Pairwise independence is not mutual independence in larger context

A_1	A_2
A_4	A_3

$P(A_1) = P(A_2) = P(A_3) = P(A_4) = 1/4$

$$A = A_1 \cup A_2; P(A)$$

$$B = A_1 \cup A_3; P(B)$$

$$C = A_1 \cup A_4; P(C)$$

$*P(ABC)$ *is the shorthand for* $P(A \cap B \cap C)$

Mutual independence

- Mutual independence of a collection of events $A_1, A_2, A_3 ... A_n$ is :

$$P(A_i|A_j A_k ... A_p) = P(A_i)$$

$$j, k, ...p \neq i$$

- It's very strong independence!

Probability using the property of Independence: Airline overbooking (1)

- An airline has a flight with 6 seats. They always sell 7 tickets for this flight. If ticket holders show up independently with probability **p**, what is the probability that the flight is overbooked ?

Probability using the property of Independence: Airline overbooking (1)

- An airline has a flight with 6 seats. They always sell 7 tickets for this flight. If ticket holders show up independently with probability **p**, what is the probability that the flight is overbooked ?

P(7 passengers showed up)

Probability using the property of Independence: Airline overbooking (2)

- An airline has a flight with 6 seats. They always sell 8 tickets for this flight. If ticket holders show up independently with probability **p**, what is the probability that exactly 6 people showed up?

P(6 people showed up) =

Probability using the property of Independence: Airline overbooking (3)

- An airline has a flight with 6 seats. They always sell 8 tickets for this flight. If ticket holders show up independently with probability **p**, what is the probability that the flight is overbooked ?

P(overbooked) =

Probability using the property of Independence: Airline overbooking (4)

- An airline has a flight with **s** seats. They always sell **t** (**t**>**s**) tickets for this flight. If ticket holders show up independently with probability **p**, what is the probability that exactly **u** people showed up?

P(exactly **u** people showed up)

Probability using the property of Independence: Airline overbooking (5)

- An airline has a flight with **s** seats. They always sell **t** (**t**>**s**) tickets for this flight. If ticket holders show up independently with probability **p**, what is the probability that the flight is overbooked ?

P(overbooked)

Independence vs Disjoint

- Q. Two disjoint events that have probability> 0 are certainly dependent to each other.
  - A. True
  - B. False

Independence of empty event

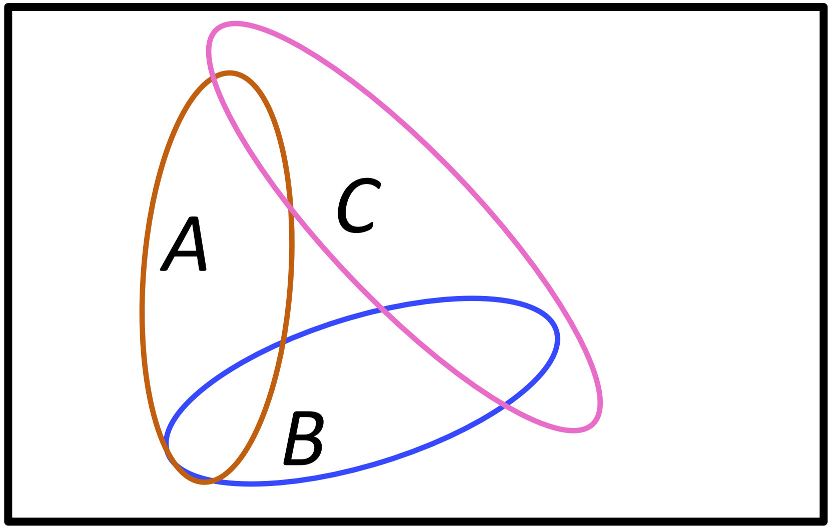
- Q. Any event is independent of empty event B.

    A. True

    B. False

Condition may affect Independence

- Assume event ***A*** and ***B*** are pairwise independent



Given ***C***, ***A*** and ***B*** are not independent any more because they become disjoint

Conditional Independence

- Event ***A*** and ***B*** are conditional independent given event ***C*** if the following is true.

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

See an example in Degroot et al. Example 2.2.10

Assignments

- Module week3 on Compass
- Next time: Random variable

Additional References

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

See you next time

See You!