COMP 1002

Logic for Computer Scientists

Lecture 28
Puzzle: Monty Hall problem

• Let’s make a deal!
  – A player picks a door.
  – Behind one door is a car.
  – Behind two others are goats.
• A player chooses a door.
  – A host opens another door
  – Shows a goat behind it.
  – And asks the player if she wants to change her choice.
• Should she switch?
Probabilities of events

- Probability of an event A is a sum of probabilities of the outcomes in A:
  \[ \Pr(A) = \sum_{a \in A} \Pr(a) \]

- Probability of A not occurring:
  \[ \Pr(\overline{A}) = 1 - \Pr(A) \]

- Probability of the union of two events (either A or B happens) is
  \[ \Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B) \]
  - By principle of inclusion-exclusion
    - If A and B are disjoint, then \( \Pr(A \cap B) = 0 \), so \( \Pr(A \cup B) = \Pr(A) + \Pr(B) \)

- In general, if events \( A_1 \ldots A_n \) are pairwise disjoint
  - that is, \( \forall i, j \text{ if } i \neq j \text{ then } A_i \cap A_j = \emptyset \)
  - Then \( \Pr(\bigcup_{i=1}^{n} A_i) = \Pr(A_1 \cup A_2 \cup \ldots \cup A_n) = \sum_{i=1}^{n} \Pr(A_i) \)
    - That is, probability of that any of the events happens is the sum of their individual probabilities.
Puzzle: Monty Hall problem

• Let’s make a deal!
  – A player picks a door.
  – Behind one door is a car.
  – Behind two others are goats.
• A player chooses a door.
  – A host opens another door
  – Shows a goat behind it.
  – And asks the player if she wants to change her choice.
• Should she switch?
  – Originally, probability of picking the car is 1/3
  – If she first picked a door with a car: (1/3 probability)
    • Then she would switch to a goat.
  – If she first picked a door with a goat (2/3 probability)
    • Then she would switch to a car.
• **Conditional probability** of an event $A$ given event $B$, denoted $\Pr(A|B)$, is the probability of $A$ if we know that $B$ occurred.
  – Probability of car a behind door 2 if we chose door 1, and door 3 had a goat behind it.

• So it is probability of both $A$ and $B$, given that we know $B$ happened for sure:

\[
\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)}
\]

  – Assume that $\Pr(B) > 0$: after all, $B$ did happen.
Independent events

• If knowing B gives us no information about A and vice versa, then A and B are **independent** events:
  
  – Then $\Pr(A) = \Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)}$.
  
  – So A and B are independent iff $\Pr(A \cap B) = \Pr(A) \cdot \Pr(B)$.

• In general, events $A_1 \ldots A_n$ can be **pairwise independent** (that is, any two $A_i, A_j$) are independent, or (stronger condition) **mutually independent**:

  \[
  \Pr(\cap_{i=1}^n A_i) = \prod_{i=1}^n \Pr(A_i)
  \]

  – For example, different coin tosses/dice rolls are usually considered independent.
Bayes theorem

- Consider a medical test that
  - Has false positive rate of 3% (healthy labeled as sick)
    - Specificity 97%
  - Has false negative rate of 1% (sick labeled as healthy).
    - Sensitivity 99%
- What is the probability that a person has the disease given that the test came positive?
  - Let A: person tested positive, B: person is sick.
    \[ \Pr(B|A) \]
  - \[ \Pr(A|B) = 0.99, \Pr(\bar{A}|B) = 0.01 \ldots \]
- Not enough information!
Bayes theorem

- **Bayes theorem** allows us to get $\Pr(B|A)$ from $\Pr(A|B)$, if we know probabilities of $A$ and $B$:

  $$\Pr(B|A) = \frac{\Pr(A|B) \cdot \Pr(B)}{\Pr(A)} = \frac{\Pr(A|B) \cdot \Pr(B)}{\Pr(A|B) \cdot \Pr(B) + \Pr(A|\bar{B}) \cdot \Pr(\bar{B})}$$

- Proof:
  - $\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)}$. $\Pr(B|A) = \frac{\Pr(A \cap B)}{\Pr(A)}$.
  - $\Pr(A \cap B) = \Pr(A|B) \Pr(B) = \Pr(B|A) \Pr(A)$
  - So $\Pr(B|A) = \Pr(A|B) \Pr(B) / \Pr(A)$

- The formula $\Pr(A) = \Pr(A|B) \Pr(B) + \Pr(A|\bar{B}) \Pr(\bar{B})$ comes from writing probability of $A$ (e.g., a positive test) as sum of probabilities of $A$ for $B$ (sick people) and for $\bar{B}$ (healthy people).
Bayes theorem

• Consider a medical test that
  – Has false positive rate of 3% (healthy labeled as sick).
  – Has false negative rate of 1% (sick labeled as healthy).
  – Tests for a disease that occurs in 5 in 1000 people.

• What is the probability that a person has the disease given that the test came positive?
  – Let A: person tested positive, B: person is sick. Pr(B|A)?
  – Pr(A|B) = 0.99, Pr(\(\bar{A}|B\)) = 0.01. Pr(\(\bar{B} | \bar{B}\)) = 0.97, Pr(A|\(\bar{B}\)) = 0.03
  – Pr(B) = 0.005.
  – Pr(A) = Pr(A|B)Pr(B) + Pr(A|\(\bar{B}\))Pr(\(\bar{B}\)) = 0.0348
  – By Bayes theorem, Pr(B|A) = \(\frac{Pr(A|B) \cdot Pr(B)}{Pr(A)}\) = 0.1422

• So the probability that a person who tested positive has the disease is just 0.1422, that is, 14.22%.
  – By a similar argument, probability that a person who tested negative does not have a disease is whopping 0.99995 = 99.995%.
• Suppose \( n \) men came to an event, and checked in their hats at the door.
  
  – On the way out, in a hurry, they each picked up a random hat.
  
  – On average, how many men picked their own hat?
Expectations

• Often we are interested in what outcome we would see “on average”.
  – How fast does this program run “on average”?

• Suppose that possible outcomes of an experiment are numbers $a_1, \ldots, a_n$
  – E.g., time a program takes to sort $n$ elements

• Its **expected value (mean)** is $\sum_{k=1}^{n} a_k \Pr(a_k)$
  – Often phrased in terms of a “random variable” $X$, where $X$ is a *function* from outcomes to numbers.
  – Write $E(X)$ to mean the expectation of $X$. 