CENG 551 Probability and Stochastic Processes for Engineers

Week # 2

Conditional Probability and Independence

Dr. Deniz Ozdemir
Deniz.ozdemir@yasar.edu.tr
Web: dozdemir.yasar.edu.tr

Plan for the Session

- Conditional Probability
- Three important tools:
  - Total probability theorem
  - Bayes' rule
  - Multiplication rule
- Independence

7 Axioms of Algebra of Events

$A \cup B = B \cup A$	*Commutative Law*
$A \cup (B \cup C) = (A \cup B) \cup C$	*Associative Law*
$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$	*Distributive Law*
$(A^c)^c = A$	
$(A \cap B)^c = A^c \cup B^c$	*DeMorgan's Law*
$A \cap A^c = \phi$	
$A \cap U = A$	

Axioms of Probability

- Sample Space: List of all possible outcomes
- Event: a subset of the sample space
- Every outcome is an event and certain set of outcomes are events
- Probability is assigned to events
- Axioms:
  1. For Any Event A, $P(A) \geq 0$
  2. $P(U) = 1$ (Normalization)
  3. If $A \cap B = \phi$, then $P(A \cup B) = P(A) + P(B)$

Axioms of Probability

- Axioms:
  - For Any Event A, $P(A) \geq 0$
  - $P(U) = 1$ (Normalization)
  - If $A \cap B = \phi$, then $P(A \cup B) = P(A) + P(B)$

From this and the prior Axioms one can determine the probability measure of an event by simply summing up all the measures for each of the finest grain events that the event consists of.

$$P(\{s_1, s_2, \ldots, s_k\}) = P(s_1) + \cdots + P(s_k)$$

- Axiom 3 needs strengthening

5

Axioms of Probability

- Axioms:
  - For Any Event A, $P(A) \geq 0$
  - $P(U) = 1$ (Normalization)
  - If $A \cap B = \phi$, then $P(A \cup B) = P(A) + P(B)$
- Exercise:
  - If $A \cap B \neq \phi$, then $P(A \cup B) = ?$

6

Example: Radar Device

- Radar device, with 3 readings:
  - Low (0), Medium (?), High (1)
- Probabilistic Modeling:
  - Sample Space / Outcomes:
- Airplane Presence + Radar Reading
  - Probability Law:

Airplane \ Radar	Low(0)	Medium(?)	High(1)
Absent	0.45	0.20	0.05
Present	0.02	0.08	0.20

Example: Radar Device

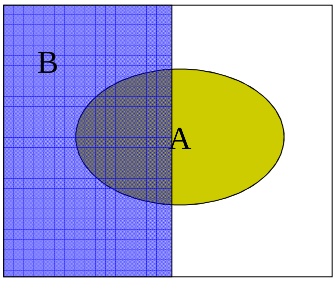
Airplane \ Radar	Low(0)	Medium(?)	High(1)
Absent	0.45	0.20	0.05
Present	0.02	0.08	0.20

- Questions:
  - What is the probability that the radar reads a medium level (?) if there are no airplanes?
  - What is the probability of having an airplane?
  - What is the probability of the airplane being there if the radar reads low (0)?
  - When should we decide there is an airplane, and when should we decide there is none?

Conditional Probability

- $P(A \mid B)$ probability of A given that B occurred.
  - B becomes our universe
- **Definition:** Assuming $P(B) \neq 0$, we have

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



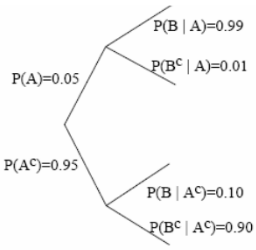
- **Consequences**: If $P(A) \neq 0$, $P(B) \neq 0$, then

$$P(A \cap B) = P(B).P(A \mid B) = P(A).P(B \mid A)$$

9

Example: Radar Device 2

- Event A: Airplane is flying above
- Event B: Something registers on radar screen



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

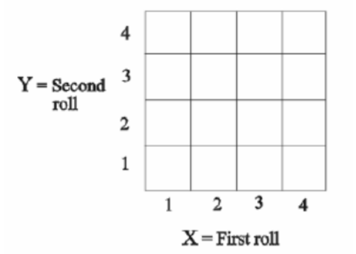
10

Example: Radar Device

Airplane \ Radar	Low(0)	Medium(?)	High(1)
Absent	0.45	0.20	0.05
Present	0.02	0.08	0.20

- Questions:
  - What is the probability that the radar reads a medium level (?) if there are no airplanes?
- Event "Absent"= Plane is absent.
  - P(Medium|Absent) =

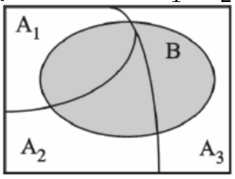
Example: Die Roll



- Let Z be the event: $min(X, Y) = 2$
- Let $M = max(X, Y)$
  - $P(M = 1 \mid Z)=$
  - $\mathrm{P}(M = 2 \mid Z)=$

Total Probability Theorem

- Divide and conquer.
- Partition of sample space into A_1, A_2, and A_3.



- One way of computing ***P****(B)*:

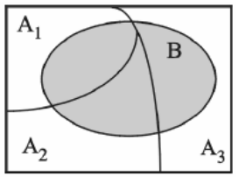
$$\begin{aligned} \mathbf{P}(B) = \quad & \mathbf{P}(A_1)\mathbf{P}(B|A_1) \\ + \ & \mathbf{P}(A_2)\mathbf{P}(B|A_2) \\ + \ & \mathbf{P}(A_3)\mathbf{P}(B|A_3) \end{aligned}$$

Value of information

- You have 100 coins. Of these 99 are normal fair coins with a probability of 0.5 landing heads and 0.5 of landing tails. One *special* coin has heads on both sides: it always comes up heads.
  - Your friend picks a coin at random. What is the probability that the coin he picked is the *special* coin?
  - Your friend then flips the coin he picked 5 times and on every flip it comes up heads. Using this additional information, what is the probability that the coin he picked is the *special* coin?

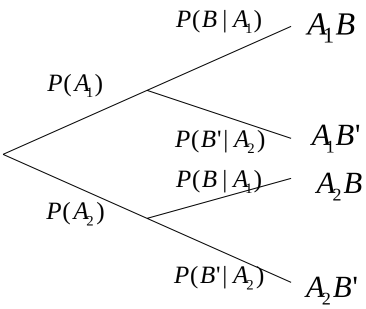
Bayes Rule

- Rules for combining evidence ("inference").
- We have "prior" probabilities: $P(A_i)$
- For each ***i***, we know: $P(B \mid A_i)$
- We wish to compute: $P(A_i \mid B)$



Bayes Theorem

- Sample Space Interpretation



$$P(A_1 \mid B) = \frac{P(A_1 B)}{P(B \mid A_1)P(A_1) + P(B \mid A_2)P(A_2)}$$

Generalized

$$P(A_i \mid B) = \frac{P(B \mid A_i)P(A_i)}{\sum_i P(B \mid A_i)P(A_i)}$$

Example: Radar Device

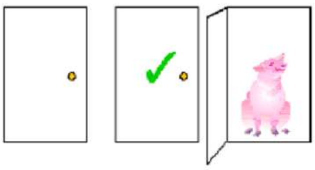
Airplane \ Radar	Low(0)	Medium(?)	High(1)
Absent	0.45	0.20	0.05
Present	0.02	0.08	0.20

- Questions:
  - What is the probability of the airplane being there if the radar reads low (0)?

- $P(\text{Present}|\text{Low}) =$

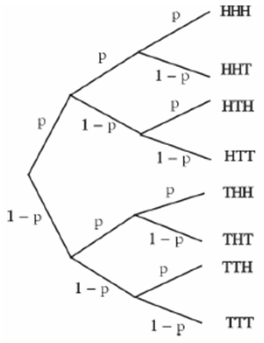
Monty Hall

- Three doors (A,B,C) behind one is your dream car, other two doors disappointment
- You select say door A. Monty, who knows where the car is, opens say door B which is empty (as he perpetrated) and offers to let you switch. What should you do?
  - Keep your door?
  - Switch the door?
  - Does not matter?



Example 2: Coin Tosses

- Look at 3 tosses of a biased coin:
  - P(Head)= p, P(Tail) =1-p



  - $P(THT)$ =
  - P(1 head)=
  - P(first toss is H|1 head) =

Radar Device: Decision Rule

- Given the radar reading, what is the best decision about the plane?
- Criterion for decision:
  - Minimize "Probability of Error"
- Decision rules:
  - Decide absent or present for each reading.
- What is the optimal decision region?

Radar / Airplane	Low(0)	Medium(?)	High(1)
Absent	0.45	0.20	0.05
Present	0.02	0.08	0.20

Radar Device: Decision Rule

- P(Error)=?
- Error={Present and decision is absent}

or {Absent and decision is present}

- Disjoint event!
- P(Error)=

Airplane \ Radar	Low(0)	Medium(?)	High(1)
Absent	0.45	0.20	0.05
Present	0.02	0.08	0.20

Exercise: HIV Testing

	HIV +	HIV -	
Test positive (+)	95	495	590
Test negative (-)	5	9405	9410
	100	9900	10000

P(HIV +) =

P(Test + | HIV +) =

P(Test -| HIV -) =

Want these to be high

P(Test -| HIV +) =

P(Test + | HIV -) =

Want these to be low

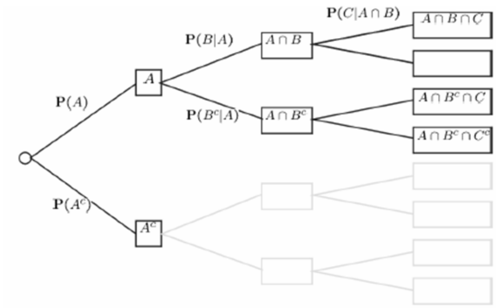
P(HIV + | Test +) =

This is one reason why we don't have mass HIV screening

Multiplication Rule

- **Example 3**: Three cards are drawn from a 52-card deck. What's the probability that none of these cards is a heart?

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



Extended Radar Example

- Threat alert affects the outcome

$P(\ldots \mid \text{No Threat})$

Airplane \ Radar	Low(0)	Medium(?)	High(1)
Absent	0.45	0.20	0.05
Present	0.02	0.08	0.20

$P(\ldots \mid \text{Threat})$

Airplane \ Radar	Low(0)	Medium(?)	High(1)
Absent	0.1125	0.05	0.0125
Present	0.055	0.22	0.55

- $P(\text{Threat})$ = Prior probability of threat = p
- What is $P(\text{airplane, radar})$?

Extended Radar Example

- A=Airplane, R=Radar Reading

$P(A, R) = P(A, R|\text{Threat}) \,.\, P(\text{Threat}) +$

$P(A, R|\text{ No Threat}) \,.\, P(\text{No Threat})$

- If we let $P(\text{Threat}) = p$, then we get:

$P(A, R)$

Radar / Airplane	Low(0)	Medium(?)	High(1)
Absent	0.45-0.3375p	0.20-0.15p	0.05-0.0375p
Present	0.02+0.0145p	0.08+0.14p	0.20+35p

Extended Radar Example

Radar / Airplane	Low(0)	Medium(?)	High(1)
Absent	0.45-0.3375p	0.20-0.15p	0.05-0.0375p
Present	0.02+0.0145p	0.08+0.14p	0.20+35p

- Given the Radar registered High, and a plane was absent, What is the probability that there was a threat?
- How does the decision region behave, as a function of p?

Exercise: Sensor Problem

- Assume that there are two chemical hazard sensors: A and B.
- Let P(A falsely detecting a hazardous chemical)=0.05 and the same for B.
- What is the probability of both sensors falsely detecting a hazardous chemical?

- If A and B are both “fooled” by the same chemical substance?

Independence of Two Events

- Definition:

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

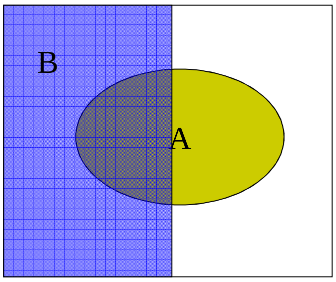
- Recall:
  - Independence of B from A: $P(B|A) = P(B)$
  - By symmetry, $P(A|B) = P(A)$
- Examples:
  - A and B are disjoint.
  - Independence of A^c and B
  - $P(A|B) = P(A|B^c)$

Independence

- Simply put $P(A|B)= P(A)$
- This implies that

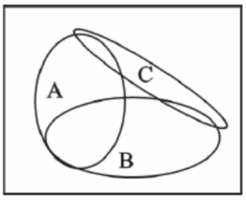
$$P(A \cap B)=P(A|B)\ P(B) = P(A)\ P(B)$$

- Interpretation in Event space:



Conditioning may affect independence

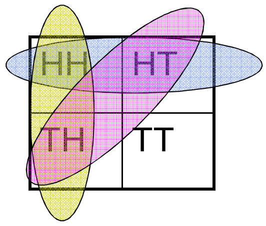
- Assume A and B are independent:



- If we are told that C occurred, are A and B independent?

Example 1

- Two independent fair (p=½) coin tosses.
  - Event A: First toss is H
  - Event B: Second toss is H
  - $\mathbf{P}(A) = \mathbf{P}(B) = 1/2$

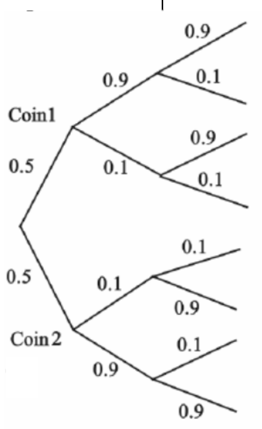


  - Event C: The two outcomes are different.
  - Conditioned on C, are A and B independent?

Example

- Choice between two unfair coins, with equal probability.
  - $P(\mathrm{H}|\mathrm{coin1}) = 0.9$
  - $P(\mathrm{H}|\mathrm{coin2}) = 0.1$
- Keep tossing the chosen coin.
- Are future tosses independent ?
  - If we know we chose coin 1?
  - If we do not know which coin we chose?
  - Compare P(toss 11 is H)

    P(toss 11 is H| first 10 tosses are H)



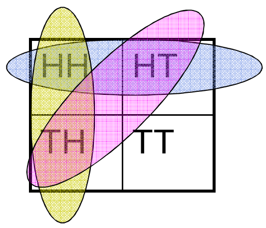
Independence of a Collection of Events

- Intuitive definition:
  - Information about some of the events tells us nothing about probabilities related to remaining events.
  - Example: $P(A_1 \cup (A_2 \cap A_3) | A_4 \cap A_5) = P(A_1 \cup (A_2 \cap A_3))$
- Mathematical definition:
  - For any distinct $i, j, \ldots, q$

    $P(A_i \cap A_j \cap \ldots \cap A_q) = P(A_i)\, P(A_i) \ldots P(A_i)$

Independence vs. Pairwise Independence

- Example 1 revisited
  - Two independent fair (p=½) coin tosses.
  - Event A: First toss is H
  - Event B: Second toss is H
  - Event C: The two outcomes are different.
  - **P**(A) = **P**(B) = **P**(C) =1/2
- P($A \cap C$) =
- P(C|A $\cap$ B) =



- Pairwise independence **does not** imply independence.