UCF Stands For Opportunity

CDA6530: Performance Models of Computers and Networks

Chapter 1: Review of Practical Probability

SCHOOL OF ELECTRICAL ENGINEERING & COMPUTER SCIENCE

Probability Definition

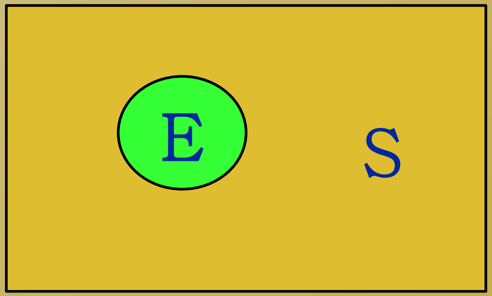
- Sample Space (S) which is a collection of objects (all possible scenarios or values). Each object is a sample point.
  - Set of all persons in a room
  - {1,2,…,6} sides of a dice
  - {0,1} for shooter results
  - (0,1) real number
- An event E is a set of sample points
  - Event $E \subseteq S$

Probability Definition

- Probability P defined on events:
  - $0 \leq P(E) \leq 1$
  - If $E=\phi$ $P(E)=0$; If $E=S$ $P(E)=1$
  - If events A and B are mutually exclusive,

    $P(A \cup B) = P(A) + P(B)$
- Classical Probability P:
  - P(E)= # of sample points in E /

    # of sample points in S



- A^c is the complement of event A:
  - A^c = {w: w not in A}
  - $P(A^c)=1-P(A)$
- Union: $A \cup B$ = {w: w in A or B or both}
- Intersection: $A \cap B$={w: in A and B}
- $P(A \cup B)=P(A)+P(B)-P(A \cap B)$
  - How to prove it based on probability definition?

  - For simplicity, define $P(AB)=P(A \cap B)$

Conditional Probability

- Meaning of P(A|B)
  - Given that event B has happened, what is the probability that event A also happens?
  - $P(A|B) = P(AB)/P(B)$
    - Physical meaning? (hint: use graph)
- Constraint sample space (scale up)

$$P(s|B) = \begin{cases} P(s)/P(B) & \text{if } s \in B \\ 0, & \text{otherwise} \end{cases}$$

Example of Conditional Probability

- A box with 5000 chips, 1000 from company X, other from Y. 10% from X is defective, 5% from Y is defective.
- A="chip is from X", B="chip is defective“
- Questions:
  - Sample space?
  - P(B) = ?
  - $P(A \cap B)$ = P(chip made by X and it is defective)
  - $P(A \cap B)$ =?
  - $P(A|B)$ = ?

  - $P(A|B)$? $P(AB)/P(B)$

Statistical Independent (S.I.)

- If A and B are S.I., then P(AB) = P(A)P(B)
  - P(A|B) = P(AB)/P(B) = P(A)
- Theory of total probability
  - $P(A) = \sum_{j=1}^{n} P(A|B_j)P(B_j)$

    where $\{B_j\}$ is a set of mutually exclusive exhaustive events, and $B_1 \cup B_2 \cup \ldots B_n = S$
  - Let's derive it for n=2:
    - $A = AB \cup AB^c$ mutually exclusive
    - $P(A) = P(AB) + P(AB^c)$

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

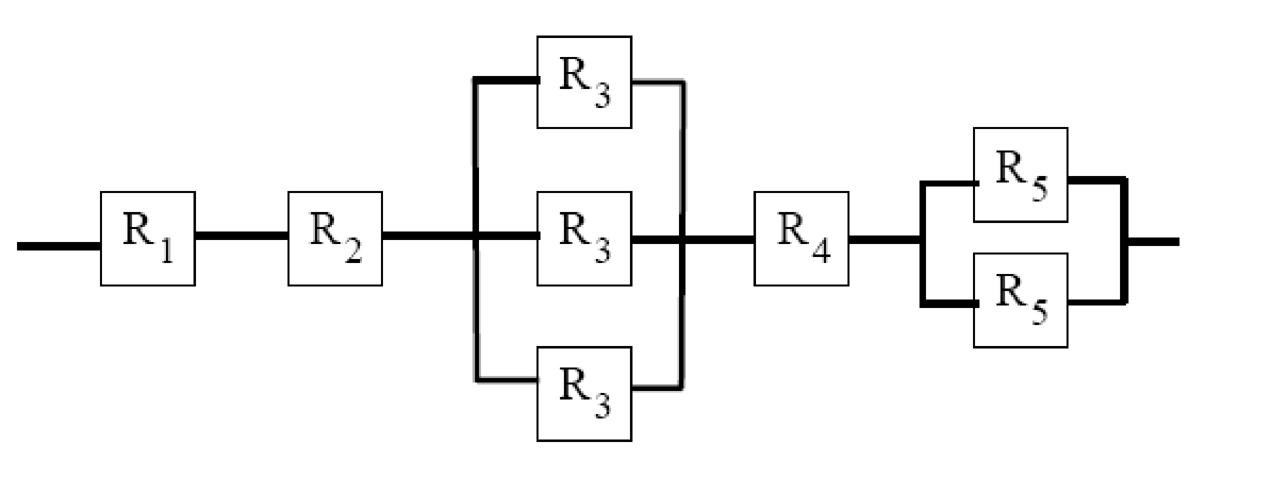
Example of Law of Total Probability

- A man shoots a target. When sunny day, he has 0.8 prob. to hit the target; when raining day, he has 0.4 prob. to hit. The weather has 0.7 prob. to be sunny, and 0.3 prob. to be raining.

- P(hit the target today)?

Application of S.I.

- R_i: reliability of component i
  - R_i = P(component i works normally)



$$R_{sys} = R_1 \cdot R_2 \cdot [1-(1-R_3)^3] \cdot R_4 \cdot [1-(1-R_5)^2]$$

Simple Derivation of Bayes' Formula

- Bayes:

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

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

- Conditional prob.:

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

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

Bayes' Theorem

- Calculate posterior prob. given observation
  - Events $\{F_1, F_2, \cdots, F_n\}$ are mutually exclusive
  - $\bigcup_{i=1}^{n} F_i = S$
  - E is an observable event
  - $P(E|F_i)$, $P(F_i)$ are known
- As E happens, which F_k is mostly likely to have happened?

$$P(F_k|E) = \frac{P(E|F_k)P(F_k)}{\sum_{i=1}^{n} P(E|F_i)P(F_i)}$$

- Law of total prob. $P(E) = \sum_{i=1}^{n} P(E|F_i)P(F_i)$

Example 1

- A man shoots a target. When sunny day, he has 0.8 prob. to hit the target; when raining day, he has 0.4 prob. to hit. The weather has 0.7 prob. to be sunny, and 0.3 prob. to be raining.

- Q: the man misses the target today, what is prob. that today is sunny? Raining?
  - The raining prob. is enlarged given the shooting result

Example 2

- A blood test is 95% accurate (detects a sick person as sick), but has 1% false positive (detects a healthy person as sick). We know 0.5% population are sick.
- Q: if a person is tested positive, what is the prob. she is really sick?
- **Model**: D: Alice is sick, E: Alice is tested positive
- Q: P(D|E)?
- Solution: It is easy to know that P(E|D) = 0.95, P(D)=0.005
- Thus we use Bayes formula
- P(D|E) = P(E|D)P(D)/P(E)
- Law of total prob.: $P(E)=P(E|D)P(D)+P(E|D^c)P(D^c)$
- =0.95*0.005+0.01*0.995
- Thus: P(D|E) = 0.323
- Testing positive only means suspicious, not really sick, although testing has only 1% false positive.
  - Worse performance when P(D) decreases.
  - Example: whether to conduct breast cancel testing in younger age?

Bayes Application ---- Naïve Bayes Classification

- Email: Spam (S) or non-spam (H)
  - From training data, we know: $P(w_i|S)$, $P(w_i|H)$
    - w_i: keyword *i* in an email
  - Define E: the set of keywords contained in an email
  - For any email, $P(E|S)=\prod P(w_i|S)$, $P(E|H)=\prod P(w_i|H)$
    - Implicit assumption that keywords are independent
  - Q: for an email, prob. to be a spam(ham)?
  - Model for Question: $P(S|E)$, $P(H|E)$

$$P(S|E) = \frac{P(E|S)P(S)}{P(E)}$$

$$P(H|E) = \frac{P(E|H)P(H)}{P(E)}$$

Reference: Naive Bayes classifier
http://en.wikipedia.org/wiki/Naive_Bayes_classifier

- Questions?