CENG 222
Statistical Methods for Computer Engineering

Spring 2021-2022

Section 1

Tolga Can (Office: B-109)

e-mail: tcan@ceng.metu.edu.tr

Online Materials will be available on ODTU-Class

Goals of the course

- Learn techniques and tools to be able to:
  - analyze and interpret large scale data,
  - apply probability theory and statistics to handle uncertainty,
  - infer facts and relationships from collected data, and
  - construct simulations by sampling from arbitrary distributions
- Acquire skills for the hot new CS field: "Data Science"

Course outline

- See the syllabus on ODTU-Class

Grading

- Midterm exam – 30%
- Final exam – 30%
- 4 Assignments –32%
- 10 Quizzes (best 8 of 10) – 8 %

ODTU Class

- Syllabus
- Course textbook
- Lecture slides and reading materials
- Assignments
- Grades

COW

- A forum for course related discussions

Textbook

- Probability and Statistics for Computer Scientists, Second Edition, Michael Baron, 2013

- Your main resource of study for this course

Probability

- Studies uncertainty

- A random experiment

  - An experiment/observation which does not have a certain outcome before it is conducted

    - Examples

      - Tossing a coin
      - Observing the life time of a light bulb
      - Number of games the Cavaliers will win this season
      - Others?

Sample space

- The set of all possible outcomes of a random experiment is called the sample space
  - Tossing a coin:
    - Sample space = {H, T}
  - Tossing two coins:
    - Sample space = {HH, HT, TH, TT}
  - Lifetime of a light bulb:
    - Sample space = $[0, +\infty)$

Event

- Any collection of possible outcomes of an experiment
  - Any subset of the sample space
- Examples:
  - Experiment: tossing two coins. Event: obtaining exactly one head. $\{HT,TH\} \subset \{HH,HT,TH,TT\}$
  - Experiment: lifetime of light bulb. Event: light bulb does not last more than a month.

    $[0,1] \subset [0,+\infty)$

Event

- A sample space of N possible outcomes yields 2^N possible events

- Example: tossing a dice once

- Sample space = {1,2,3,4,5,6}

- Number of possible events = $2^6 = 64$

- Example events?

Notation used in the book

- Ω = sample space

- $\varnothing$ = empty event

- $P\{E\}$ = probability of event E

Event algebra

- Union of two events: same as set union
  - $A \cup B = \{x: x \in A \text{ or } x \in B\}$
- Intersection of two events: same as set intersection
  - $A \cap B = \{x: x \in A \text{ and } x \in B\}$
- Complementation: same as in sets
  - A^c or $\bar{A} = \{x: x \in \Omega \text{ and } x \notin A\}$
- Difference: same as in sets
  - $A \backslash B = \{x: x \in A \text{ and } x \notin B\}$

Disjoint and exhaustive events

- Disjoint events: If A and B have no outcomes in common, i.e., $A \cap B = \varnothing$
  - Also called mutually exclusive events
- If the union of a number of events equals the sample space, they are called exhaustive
  - $A \cup B \cup C = \Omega$

Complement, Union, Intersection

- $\overline{A \cup B} = \bar{A} \cap \bar{B}$

- $\overline{A \cap B} = \bar{A} \cup \bar{B}$

- $\overline{E_1 \cup E_2 \cup E_3 \cup E_4} = \overline{E_1} \cap \overline{E_2} \cap \overline{E_3} \cap \overline{E_4}$

- $\overline{E_1 \cap E_2 \cap E_3 \cap E_4} = \overline{E_1} \cup \overline{E_2} \cup \overline{E_3} \cup \overline{E_4}$

Probability

- Assignment of a real number to an event
  - The relative frequency of occurrence of an event in a large number of experiments
- $P(A)$
- Axioms of probability:
  - $P(A) \geq 0$
  - $P(\Omega) = 1$
  - If A and B are mutually exclusive events, then $P(A \cup B) = P(A) + P(B)$
- Any function that satisfies these axioms is called a probability function

Example

- Experiment:
  - Tossing two coins
  - $A = $ {obtaining exactly one head}
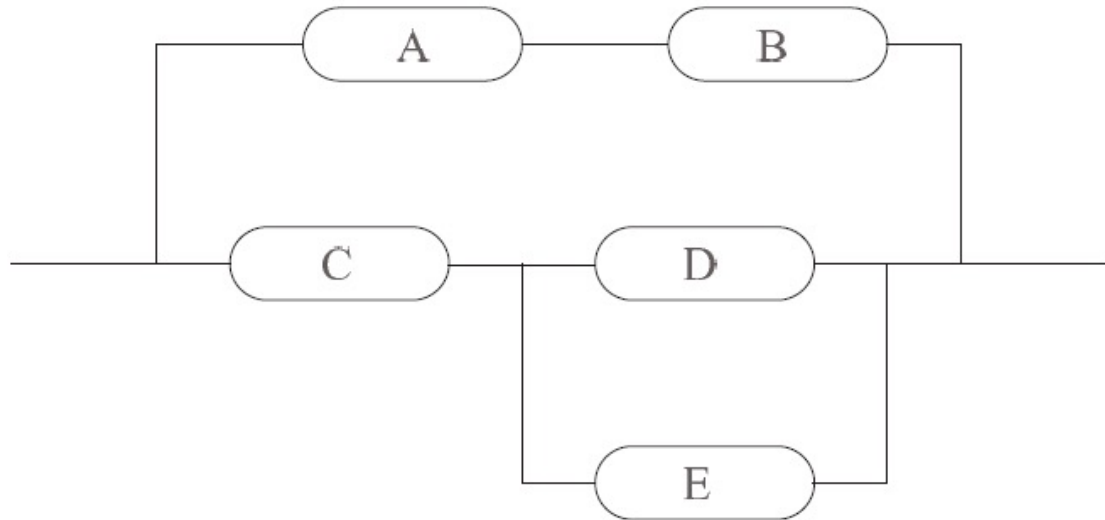 - $P(A) = $?

Computing probabilities

- for non-"mutually exclusive" events:
  - $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Independent Events

- $P(E_1 \cap E_2 \cap E_3) = P\{E_1\} \cdot P\{E_2\} \cdot P\{E_3\}$

Applications in reliability

- Example 2.18
- Example 2.19
- Example 2.20

Conditional probability

- Updating of the sample space based on new information

- Consider two events A and B. Suppose that the event B has occurred. This information will change the probability of event A.

- $P(A|B)$ denotes the conditional probability of event A given that B has occurred.

Conditional probability

- If A and B are events in Ω and $P(B)>0$, then $P(A|B)$ is called the conditional probability of A given B if the following axiom is satisfied:
  - $P(A|B) = P(A \cap B)/P(B)$

- Example: tossing a fair dice.
  - $A = \{$the number on the dice is even$\}$
  - $B = \{$the number on the dice $< 4\}$
  - $P(A|B) = ?$

Independence

- If $P(A|B)=P(A)$ we call that event A is independent of event B

- Note:
  - if two events A and B are independent, then

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

- Show that $P(B|A)=P(B)$ also holds in this case.
  - In other words, A and B are mutually independent

- This does NOT mean that they are disjoint. If A and B are disjoint then $P(B|A)=0$

Independence

- Example: tossing a fair dice.
    - A = {the number on the dice is even}
    - B = {the number on the dice > 2}
    - P(A|B) = ?
    - P(B|A) = ?
    - P(A) = ?
    - P(B) = ?
- Example 2.31

Bayes' Rule

- Using conditional probability formula we may write:
  - $P(A|B) = P(A \cap B)/P(B)$
  - $P(B|A) = P(A \cap B)/P(A)$
  - $\rightarrow P(A \cap B) = P(A|B)P(B) = P(B|A)P(A) \rightarrow$

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

- This is known as the Bayes' rule
- It forms the basis of Bayesian statistics
- What additional probabilities do we need to know to solve Example 2.32?

Law of Total Probability

- Let B_1, B_2, B_3, ….,B_k be a partition of the sample space. B_is are mutually disjoint. Let A be any event.

- Note that B_is also partition A

- Then for each $i = 1,2,…,k$

$$P(B_i \mid A) = \frac{P(A \mid B_i)P(B_i)}{P(A)} = \frac{P(A \mid B_i)P(B_i)}{\sum_{j=1}^{k} P(A \mid B_j)P(B_j)}$$

When $P(A)$ is not directly known, but known conditionally, we make use of this law.

Bayes' Rule for two events

$$P(B \mid A) = \frac{P(A \mid B)P(B)}{P(A)} = \frac{P(A \mid B)P(B)}{P(A \mid B)P(B) + P(A \mid \overline{B})P(\overline{B})}$$

- Now, solve Exercise 2.32, given **P**(B)

Another example

- A novel disease diagnostic kit is 95% effective in detecting a certain disease when it is present. The test also has a 1% false positive rate. If 0.5% of the population has the disease, what is the probability a person with a positive test result actually has the disease?

Solution

- $A = \{$a person's test result is positive$\}$
- $B = \{$a person has the disease$\}$
- $P(B) = 0.005$, $P(A|B) = 0.95$, $P(A|B^c) = 0.01$

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

$$= \frac{0.95 \times 0.005}{0.95 \times 0.005 + 0.01 \times (1 - 0.005)} = \frac{475}{1470} \approx 0.323$$

Random Variables

- A random variable (r.v.) associates a unique numerical value with each outcome in the sample space. It is a real-valued function from a sample space Ω into real numbers.

- Similar to events it is denoted by an uppercase letter (e.g., *X* or *Y*) and a particular value taken by a r.v. is denoted by the corresponding lowercase letter (e.g., *x* or *y*).

Examples

- Toss three coins. X = number of heads
- Pick a student from the Computer Engineering Department.

  X = age of the student
- Observe lifetime of a light bulb

  X = lifetime in minutes
- X may be discrete or continuous