ECE 2521: Analysis of Stochastic Processes

Lecture 1

Department of Electrical and Computer Engineering
University of Pittsburgh

September, 1^{st} *2021*

Azime Can-Cimino

Instructor

- Instructor: Azime Can-Cimino (azime.cancimino@pitt.edu)
    - Lectures on Wednesday: 6:00-8:30 PM, G37 Benedum Hall (**except first two weeks online via Zoom**)
    - Office Hours: Friday 10:00-11:00 AM Via Zoom (link on Syllabus)

Textbook: Probability, Statistics and Random Processes for Electrical Engineering by Alberto Leon-Garcia (3rd Edition)

Grading

- Midterm Exam: 50%
- Final Exam: 50%

- Homework: $5 - 7$ assignments. No submission!
- Exams: are mandatory! If you cannot make it, please let me know at least a week prior to the scheduled exam.

Application Areas

- Engineering
  - Communication
  - Radar systems
  - Signal processing
  - Control systems
  - Decision and resource allocation
  - Reliability
- Economics and finance
- Meteorology
- Natural sciences: physics, statistical mechanics
- Statistics (collection and organization of data so that useful inference can be drawn from them)

Mathematical Models

- **Deterministic Models**: set of mathematical equations specifies the exact outcome
- **Probability Models**: outcome varies in unpredictable fashion (random) when repeated under the same conditions
  - Define the random experiment
  - Specify the set of all possible outcomes
  - Specify the probability assignment from which the probabilities of all events of interest can be computed

Probability

- What is a probability?
  - Number between 0 and 1 inclusive that reflects the likelihood of occuring of a physical event
- Relative Frequency approach
  - Repeat experiment n times under identical conditions
  - Compute the relative frequency of the outcome $k = 0, 1, 2, \ldots$

$$f_k(n) = \frac{N_k(n)}{n}$$

  - Estimate the probability of the outcome k: $p_k = \lim_{n \to \infty} f_k(n)$
- Axiomatic approach
  - Unified mathematical theory of probability that is not tied to any particular application
  - Allows the interpretation of the probability as relative frequency
  - Set of axioms for the probability assignment to satisfy

Review of Set Theory

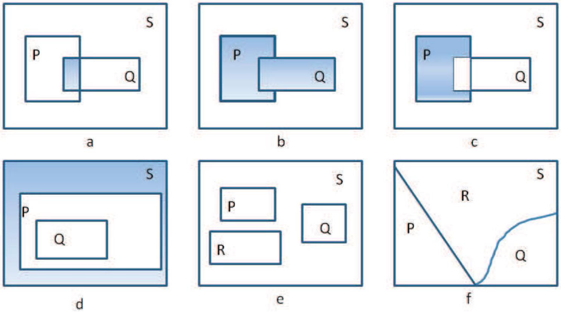
- **Set**: Collection of distinct objects.
- **Universal set**: Set containing all elements in all sets under consideration in a particular problem, denoted by S.
- **Null set**: Set that has no elements (empty), denoted by ϕ.
- Examples:
  - Define a set by listing all of its elements: $A = \{1, 4, 9, 16, 25\}$
  - Define a set by showing a rule to generate all elements of the set: $B = \{x^2 | x = 1, 2, 3, 4, 5\}$
  - Some sets can have infinite dimensions:
    - C=$\{$all positive integers$\}$
    - D=$\{$all even positive integers$\}$
    - $E = \{x^2 | x = 1, 2, 3, ...\}$

Set Operations and Properties

- **Set Theoretic Operations**:
  - Union: $\xi \in A \cup B$ iff $\xi \in A$ or $\xi \in B$.
  - Intersection: $\xi \in A \cap B$ iff $\xi \in A$ and $\xi \in B$.
  - Complement: $\xi \in A^c \Leftrightarrow$ (iff) $\xi \notin A$.
  - Difference: $\xi \in A - B \Leftrightarrow \xi \in A$ and $\xi \notin B$ ($A - B = A \cap B^c$).
  - Set equality: Any two sets A and B are equal to each other, $A = B$, if and only if (iff) $B \subset A$ and $A \subset B$.
- **Mutually Exclusive**: Two or more sets, A_1, A_2, A_3, ... are mutually exclusive (disjoint) if they have no common elements: $A_i \cap A_j = \phi, \quad$ for all $i \neq j$
- **Collectively Exhaustive**: A collection of sets $A_1, \ldots, A_n$ is collectively exhaustive iff: $\bigcup_{i=1}^{n} A_i = A_1 \cup A_2 \cup \cdots \cup A_n = S$
- **Partition**: A collection of mutually exclusive and collectively exhaustive sets form a partition of the sample space, S.

Venn Diagram

- Illustrates the relationship among sets



(a) $P \cap Q$. (b) $P \cup Q$. (c) $P \cap Q^c$. (d) $Q \subset P$. Complement of P. (e) Sets P, Q, and R are disjoint. (f) Sets P, Q, and R form a partition of the sample space S.

Elementary Set Relations

(1) Commutativity:
$$A \cup B = B \cup A$$
$$A \cap B = B \cap A$$

(2) Associativity:
$$A \cup (B \cup C) = (A \cup B) \cup C = (A \cup C) \cup B$$
$$A \cap (B \cap C) = (A \cap B) \cap C = (A \cap C) \cap B$$

(3) Distributivity:
$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$
$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

(4) $(A^c)^c = A$,

(5) $A \cap A^c = \phi$

Elementary Set Relations

(6) De Morgan's Laws

$(A \cup B)^c = A^c \cap B^c$

$(A \cap B)^c = A^c \cup B^c$

(7) $A \cap S = A$

(8) Using Commutativity and Associativity we can define

(i) $\bigcap_{i=1}^{N} A_i = A_1 \bigcap A_2 \bigcap A_3 \ldots \bigcap A_N$

(ii) $\bigcup_{i=1}^{N} A_i = A_1 \bigcup A_2 \bigcup A_3 \ldots \bigcup A_N$

(9) Countably Infinite Set Operations

(i) $\xi \in \bigcap_{i=1}^{\infty} A_i \triangleq \{\xi \in S | \xi \in A_i \text{ for all } i\}$

(ii) $\xi \in \bigcup_{i=1}^{\infty} A_i \triangleq \{\xi \in S | \xi \in A_i \text{ for at least one } i\}$

Fields and σ-Fields of Sets

- A set of subsets $\mathcal{F}$ of a universal set S is a **field** if
  (1) $S \in \mathcal{F}$
  (2) if $A \in \mathcal{F} \to A^c \in \mathcal{F}$ ($\mathcal{F}$ is closed under complement)
  (3) if $A, B \in \mathcal{F} \to A \cup B \in \mathcal{F}$ ($\mathcal{F}$ is closed under *finite* unions)
  - *Example*: If $S = \{0, 1\}$, then $\mathcal{F} = \{S, \{1\}, \{0\}, \phi\}$ is a field.
    Proposition 1: If A, $B \in \mathcal{F}$ then $A \bigcap B \in \mathcal{F}$.
    Proposition 2: If $A_i \in \mathcal{F}$ for $i = 1, \ldots, N$, then $\bigcup_{i=1}^{N} A_i \in \mathcal{F}$.
- A set of subsets of S is a σ-**field** $\mathcal{F}$ if
  (1) $S \in \mathcal{F}$
  (2) if $A \in \mathcal{F} \to A^c \in \mathcal{F}$ ($\mathcal{F}$ is closed under complement)
  (3) if $A_i \in \mathcal{F}$ for $\forall i \to \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$ ($\mathcal{F}$ is closed under *countable* unions)
  Proposition: If $A_i \in \mathcal{F}$ for $i = 1, \ldots, N$, then $\bigcap_{i=1}^{\infty} A_i \in \mathcal{F}$.
  (Hint: countable generalization of De Morgan's Law)

Probability Model

- **Experiment**: Some process that results in an outcome that is not fully predictable.
  Examples: (a) Flipping a coin. (b) Measuring the temperature in the park a week from now. (c) Signal received at the output of a noisy communication channel.
- **Outcome**: Any possible observation of the experiment.
- **Sample Space** (S): Set of all the finest grain, mutually exclusive, collectively exhaustive outcomes of the experiment.
- **Event**: An event is a set of outcomes of an experiment.
- **Event Space**: An event space is a collectively exhaustive, mutually exclusive set of events.
- **Probability Measure**: Assigns a number $P(A)$ to every event A satisfying a set of axioms.

Axioms of Probability

- Let $\mathcal{F}$ be a σ-field of a sample space S (specified for a random experiment). A function $P : \mathcal{F} \to [0, 1]$ is a **probability measure** if
  1. **Axiom I**: $P(S) = 1$
  2. **Axiom II**: $P(E) \geq 0$ for every event $E \in \mathcal{F}$
  3. **Axiom III**: If $E_i \in \mathcal{F}$ for $i = 1, 2, 3, \ldots$ are disjoint (mutually exclusive) then $P\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i)$.

- The triple $(S, \mathcal{F}, P)$ defines a **probability space**

Observations on Probability Measure

- Sample space could be continuous (S uncountable), or discrete (S finite or countable)
- When S is uncountable, for example if $S = \mathbb{R}$ or $S = \mathbb{R}^2$, σ-field (i.e. set of events of interest) is defined by countable unions, intersections, and complements of intervals in $\mathbb{R}$ or rectangles in $\mathbb{R}^2$
- For **discrete models** the probability measure is specified by the probabilities of every single event
- For **continuous models** the probabilities are determined by sub-intervals or lengths (1D), or areas (2D), or volumes (3D)

Examples

- Example 1

Experiment	Select a ball from an urn containing balls numbered 1 to 50
Sample space	$S =$
Probability	Assume outcomes are equally likely
Event	"An even ball is selected"
Determine	$P(A) =$

- Example 2

Experiment	Toss a coin three times and note the sequence of H and T
Sample space	$S =$
Probability	Assume fair coin: H and T are equally likely
Event	"The three tosees give the same outcome"
Determine	$P(A) =$

Examples

- Example 1

Experiment	Select a ball from an urn containing balls numbered 1 to 50
Sample space	$S = \{1, 2, \ldots, 50\}$
Probability	Assume outcomes are equally likely
Event	"An even ball is selected": $A = \{2, 4, \ldots, 48, 50\}$
Determine	$P(A) = 0.5$

- Example 2

Experiment	Toss a coin three times and note the sequence of H and T
Sample space	$S = \{HTT, THT, TTH, HHT, HTH, THH, HHH, TTT\}$
Probability	Assume fair coin: H and T are equally likely
Event	"The three tosees give the same outcome": $A = \{HHH, TTT\}$
Determine	$P(A) = \frac{2}{8} = 0.25$

Examples

- Example 3

Experiment Alice takes the bus to work each morning and her travel time T varies from 20 to 40 mins.

Sample space $S =$

Probability Assume all travel times in the interval are equally likely: $P(S) = P(20 \leq T \leq 40) = 1$

Determine
(a) Alice's travel time is more than 30 mins.
(b) Alice's travel time is exactly 22.01 mins.

Examples

- Example 3

Experiment Alice takes the bus to work each morning and her travel time T varies from 20 to 40 mins.

Sample space $S = \{T | 20 \leq T \leq 40 \text{ mins}\}$

Probability Assume all travel times in the interval are equally likely: $P(S) = P(20 \leq T \leq 40) = 1$

Determine

(a) Alice's travel time is more than 30 mins.
$P(T \geq 30) = 1/2$

(b) Alice's travel time is exactly 22.01 mins.
$P(T = 22.01) = 0$

Properties of Probability

- For every probabilistic model and events A, B and C:
  - **Corollary 1**: $P(A^c) = 1 - P(A)$.
  - **Corollary 2**: $P(A) \leq 1$.
  - **Corollary 3**: $P(\phi) = 0$
  - **Corollary 4**:
    $P(A \cup B \cup C) = P(A) + P(A^c \cap B) + P(A^c \cap B^c \cap C)$
  - **Corollary 5**: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
  - **Corollary 6**: $P(A \cup B) \leq P(A) + P(B)$
  - **Corollary 7**: If $A \subset B$, $P(A) \leq P(B)$

Conditional Probability

- Given a probability space $(S, \mathcal{F}, P)$ and two events $A, B \in \mathcal{F}$, where $P(B) > 0$, the conditional probability is defined as

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

- If we define $\mathcal{F}_B \triangleq \{Q \in \mathcal{F} | Q \subseteq B\}$, then $(B, \mathcal{F}_B, P(\cdot|B))$ is the new probability space
- Allows us to reason about the likelihood of an event based on partial knowledge (a priori information)
- It forms the basis for inference theory: observe the effect and reason about the cause

Conditional Probability

- All probabilities are concentrated on B because $P(B|B) = 1$
- If A and B are mutually exclusive, then

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(\phi)}{P(B)} = 0$$

- If $B \subset A$, then $P(A|B) = \frac{P(B)}{P(B)} = 1$
- If $A \subset B$, then $P(A|B) = \frac{P(A)}{P(B)} \leq 1$
- Nonnegativity: $P(A|B) = \frac{P(A \cap B)}{P(B)} \geq 0$
- Additivity: For any two disjoint (mutually exclusive) events A_1 and A_2 we have that: $P(A_1 \cup A_2|B) = P(A_1|B) + P(A_2|B)$.

Examples

Experiment Test two integrated circuits from the same silicon wafer and observe whether a circuit is accepted (a) or rejected (r).

Sample space $S = \{aa, ar, ra, rr\}$

Probabilities $P(rr) = 0.01$, $P(ra) = 0.01$, $P(ar) = 0.01$, $P(aa) = 0.97$.

Event A The second circuit is rejected $A = \{rr, ar\}$

Event B The first circuit is rejected, $B = \{rr, ra\}$

Determine $P(A)$, $P(B)$, and probability of second circuit to be rejected observing that the first one is rejected, $P(A|B)$

Solution
$$P(A) = P(rr) + P(ar) = 0.02$$
$$P(B) = P(rr) + P(ra) = 0.02$$
$$P(A \cap B) = P(\text{both rejected}) = 0.01$$
$$P(A|B) = \frac{P(A\cap B)}{P(B)} = 0.01/0.02 = 0.5$$

- The probability of rejecting the second circuit increases from 0.02 to 0.5 once we know that first circuit is rejected.

Multiplication Rule

- Provides a means to compute the occurrence of multiple events in an experiment as a result of multiple subexperiments
- The output of each subexperiment may depend on the results of previous experiments
- If $A_1, \ldots, A_n$ are events,

$$\begin{aligned} P\left(\bigcap_{i=1}^{n} A_i\right) &= P\left(A_1 \cap A_2 \cap \cdots \cap A_n\right) \\ &= P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2)\cdots P\left(A_n \middle| \bigcap_{i=1}^{n-1} A_i\right). \end{aligned}$$

Sequential Tree

- Sequential calculation of probabilities
- Construct a sequential description of the sample space.
- Record the corresponding conditional probabilities along branches of the tree.
- Obtain the probability of any one outcome using the multiplication rule.

Example

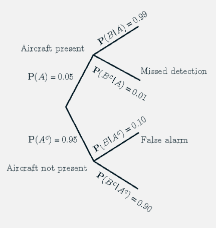
If an aircraft is present, a radar correctly registers its presence with probability 0.99. If an aircraft is absent, the radar falsely registers a presence with probability 0.1. An aircraft is present with probability 0.05.

(a) What is the probability of a false alarm (false detection of an aircraft's presence)?

(b) What is the probability of a missed detection (an aircraft is present but does not register)?

(c) What is the probability that an aircraft is present given a registration on the radar screen?

Example



Let's define the following events:

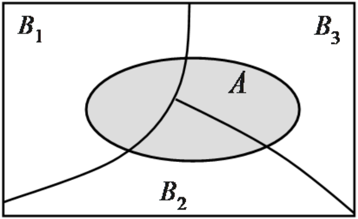
A : aircraft present, B : a registration on the radar screen

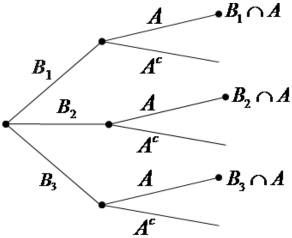
$P(\text{false alarm}) = P(A^c \cap B) = P(A^c)P(B|A^c) = (0.95)(0.1) = 0.095$

$P(\text{missed detection}) = P(A \cap B^c) = P(A)P(B^c|A) = (0.05)(0.01) = 0.0005$

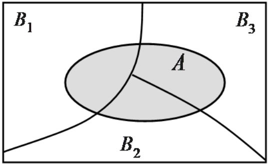
Law of Total Probability

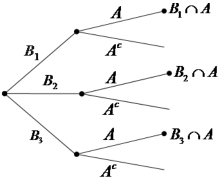
- The events $B_1, \ldots, B_n$ form a partition of the sample space such that $B = \{B_1, B_2, \ldots, B_n\}$ is an event space with B_i's mutually exclusive ($B_i \cap B_j = \phi$ for $i \neq j$) and collectively exhaustive ($B = S$)





Law of Total Probability





- The conditional probabilities for an event A, i.e. $P(A|B_1)$, ..., $P(A|B_n)$, are known, and the unconditional probability for event A using a divide and conquer approach can be calculated using:

$$\begin{aligned} P(A) &= P(A \cap B_1) + \cdots P(A \cap B_n) \\ &= P(A|B_1)P(B_1) + \cdots + P(A|B_n)P(B_n). \end{aligned}$$

Bayes' Rule

- Useful for problems where the prior probabilities $P(B_i)$ and $P(A|B_i)$ are given, $B = \{B_1, \ldots, B_n\}$ is an event space, and we would like to compute $P(B_i|A)$ when A is observed.

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

- From the total probability theorem (since B is an event space),

$$P(A) = \sum_{i=1}^{n} P(A|B_i)P(B_i)$$

- Objective is inference rather than prediction

Example: Communication Channels

A data communication system has 3 types of channels with varying quality of transmission.
The probability of receiving a bit in error is 0.01, 0.005 and 0.001 for channel types I, II and III respectively.
The three channel types are used to transfer bits in the ratio 20%, 30%, and 50% respectively.
Find the probability that a bit is received in error.
Let E denote the event that a bit is received in error. We are given, $P(E|\text{Type I}) = 0.01$ $P(E|\text{Type II}) = 0.005$ $P(E|\text{Type III}) = 0.001$, and also, $P(\text{Type I}) = 0.2$ $P(\text{Type II}) = 0.3$ $P(\text{Type III}) = 0.5$
Applying the total probability theorem, the unconditional probability for receiving an arbitrary bit in error is,

$$\begin{aligned}
P(E) &= P(E \cap \text{Type I}) + P(E \cap \text{Type II}) + P(E \cap \text{Type III}) \\
&= P(E|\text{Type I})P(\text{Type I}) + P(E|\text{Type II})P(\text{Type II}) + P(E|\text{Type III})P(\text{Type III}) \\
&= (0.01)(0.2) + (0.005)(0.3) + (0.001)(0.5) \\
&= 0.004
\end{aligned}$$

Independence of two events

- Two events A and B are independent if and only if

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

- When events A and B have non-zero probabilities, the definition of independence is equivalent to:
  $P(A|B) = P(A) \qquad P(B|A) = P(B)$
- When two events are independent, knowledge of any one of the events occurring does not alter our information (probability) of the other event.
- If events A and B are independent, their complements are also independent: $P(A^c B^c) = P(A^c)P(B^c)$.
- If events A and B are disjoint, they are neccesarily *dependent*. This is because: $P(AB) = P(\phi) = 0 \neq P(A)P(B)$.

Independence of a collection of events

- The events $A_1, \ldots, A_n$ are mutually independent if

$$P\Big(\bigcap_{i\in I} A_i\Big) = \prod_{i\in I} P(A_i)$$

for every subset I of $\{1, 2, \cdots, n\}$

Example

- Consider an experiment consisting of two independent fair coin tosses. We define the following events.
- A: First toss is a head
- B: Second toss is a head
- C: First and second tosses have different outcomes.
- Are events A, B and C mutually independent?