CIE 327 - Probability and Stochastic Processes

SalahDin A. Rezk

Zewail City of Science and Technology

s-salahdin.rezk@zewailcity.edu.eg

Student's notes for the Communication and Information Engineering Probability Course by Prof. Samy Soliman at Zewail City of Science and Technology in the Fall 24/25 semester. The notes are based on the lectures and the textbook *Probability, Random Variables, and Stochastic Processes* by Athanasios Papoulis and S. Unnikrishna Pillai. Other references include *Probability and Statistics for Engineers and Scientists* by Ronald E. Walpole and Raymond H. Myers; and *Modern Digital and Analog Communication Systems* by B. P. Lathi.and Zhi Ding. The notes cover the basics of probability theory, random variables, and stochastic processes.

Contents

List of Figures

List of Theorems

PART

I

Probability Theory

Section 1

Introduction

Probability applications are everywhere, from weather forecasting to aerospace engineering. It is a mathematical tool to model uncertainty.

Definition 1 (Random Experiment) An experiment with an uncertain outcome.

Definition 2 (Sample Space) The set of all possible outcomes.

Example (Heads or Tails) The sample space is $\{H, T\}$.

Example (Rolling a Die) The sample space is $\{1, 2, 3, 4, 5, 6\}$.

Example (Point on a Circle) The sample space is $S = \{(x, y) \mid x^2 + y^2 \leq 5\}$.

Remark The number of elements in a sample space may be finite, infinite, countable, or uncountable.

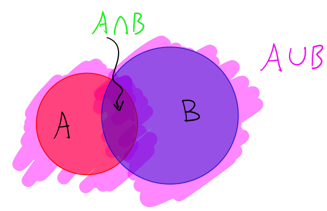

Figure 1. Example of Unions and Intersections on events A and B.

Definition 3 (Event) A subset of the sample space.

Definition 4 (Complement) The event that is not in A, denoted by A'.

Definition 5 (Union) An event that is in A or B, denoted by $A \cup B$.

Definition 6 (Intersection) An event common to both A and B, denoted by $A \cap B$.

Definition 7 (Mutually Exclusive/Disjoint) Two events are mutually exclusive/disjoint if they have no common outcomes, i.e., $A \cap B = \emptyset$.

Definition 8 (Venn Diagram) A diagram that shows the relationships between events. See Figures.

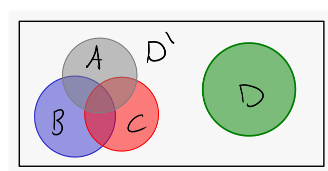

Figure 2. Venn Diagram for Sample Space S

Subsection 1.1

Some Properties of Events

$$A \cap \emptyset = \emptyset \tag{1.1}$$
$$A \cup \emptyset = A \tag{1.2}$$
$$A \cap A' = \emptyset \tag{1.3}$$
$$A \cup A' = S \tag{1.4}$$
$$S' = \emptyset \tag{1.5}$$

$$(A')' = A \tag{1.6}$$
$$(A \cup B)' = A' \cap B' \tag{1.7}$$
$$(A \cap B)' = A' \cup B'. \tag{1.8}$$

Check the Venn Diagram in Figure 3 for a visual representation of these properties.

PROOF (Equation 1.7)

$$\begin{aligned} x \in (A \cup B)' &\iff x \notin A \cup B \\ &\iff x \notin A \text{ and } x \notin B \\ &\iff x \in A' \text{ and } x \in B' \\ &\iff x \in A' \cap B'. \end{aligned}$$

□

PROOF (Equation 1.8)

$$\begin{aligned} x \in (A \cap B)' &\iff x \notin A \cap B \\ &\iff x \notin A \text{ or } x \notin B \\ &\iff x \in A' \text{ or } x \in B' \\ &\iff x \in A' \cup B'. \end{aligned}$$

□

SUBSECTION 1.2
Axioms of Probability

- $P(A) \in [0, 1]$ for all events A.
- $P(S) = 1$.
- If $A_1, A_2, \ldots$ are mutually exclusive, then $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$.

SUBSECTION 1.3
More Rules

$$P(A') = 1 - P(A) \tag{1.9}$$
$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \tag{1.10}$$
$$P(\varnothing) = 0 \tag{1.11}$$

PROOF (Equation 1.9)

$$P(A \cup A') = P(S) = 1$$

Since A and A' are mutually exclusive (Definition 7),

$$\begin{aligned} P(A) + P(A') &= 1 \\ P(A') &= 1 - P(A). \end{aligned}$$

□

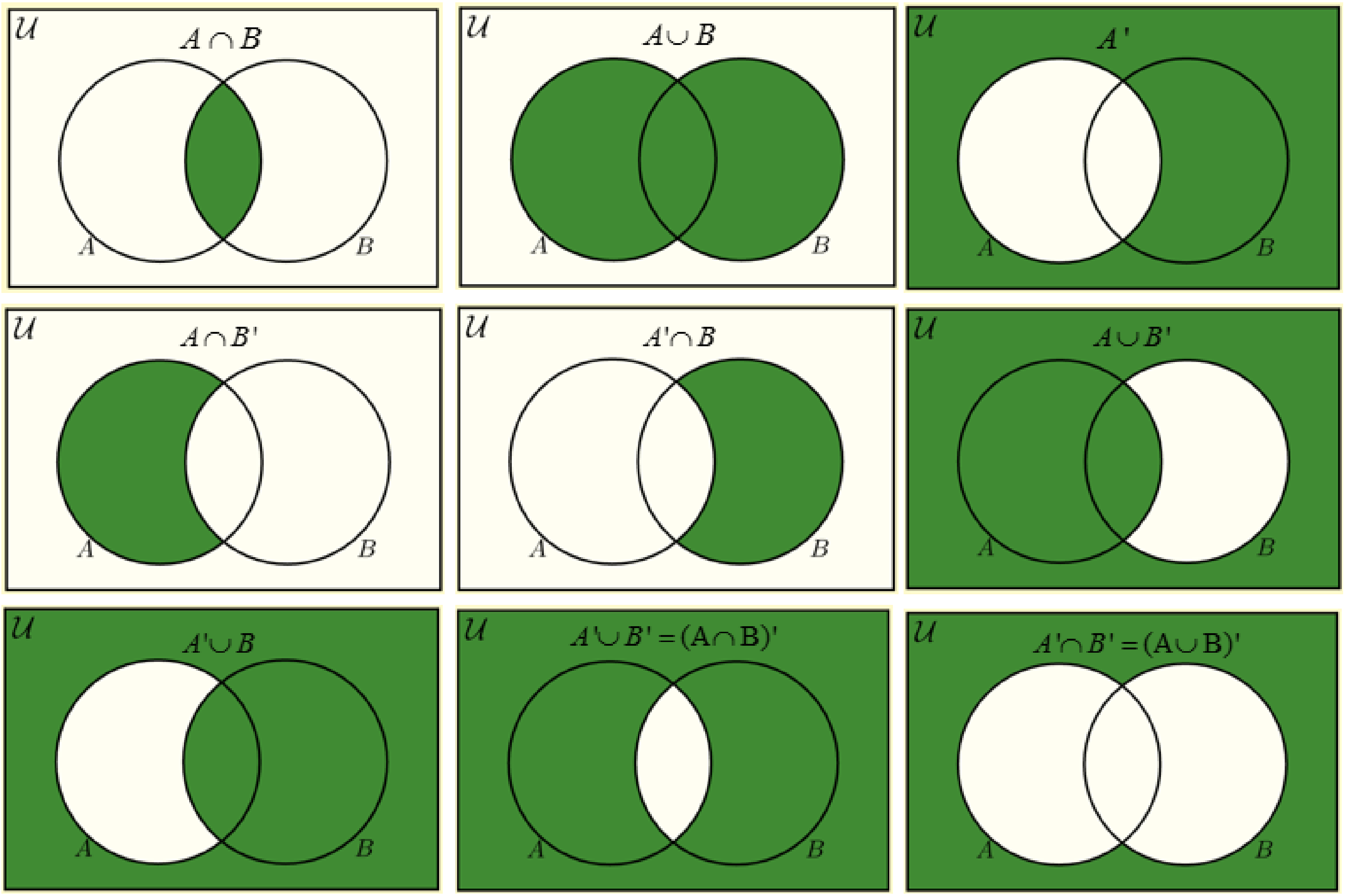

Figure 3. Venn Diagram for some proprieties.

SECTION 2

Counting Techniques

They are used to determine the number of outcomes in a sample space. Helpful for calculating probabilities and will still be useful in later chapters.

SUBSECTION 2.1

Multiplication Rule

Definition 9 (Multiplication Rule) If an experiment consists of n_1 stages, where the first stage can result in n_1 outcomes, the second stage in n_2 outcomes, and so on, the total number of outcomes is $n_1 \times n_2 \times \cdots \times n_k = \prod_{i=1}^{k} n_i$.

Example (License Plates) A license plate consists of 3 letters followed by 3 digits. Total number of plates is $26^3 \times 10^3$.

Definition 10 (Factorial) The product of all positive integers up to n.

$$n! = n \times (n-1) \times \cdots \times 1.$$

Remark

- $0! = 1$.
- $n! = n \times (n-1)!$.

Factorials are used to calculate permutations and combinations, representing the number of ways to arrange a set of objects.

Example (Arranging Professors) Nine professors are to give talks at a conference, grouped by nationality (3 French, 2 American, 4 Egyptian). In how many ways can their talks be scheduled so that professors of the same nationality follow each other?

$$\begin{aligned}
\text{Arrange French: } & 3! \\
\text{Arrange American: } & 2! \\
\text{Arrange Egyptian: } & 4! \\
\text{Arrange Nationalities: } & 3! \\
\text{Total: } & 3! \times (3! \times 2! \times 4!) = 1728.
\end{aligned}$$

Subsection 2.2

Combinatorics

Definition 11 (Permutation) An arrangement of n objects in a specific order.

$$P(n, r) = \frac{n!}{(n-r)!}.$$

Proof (Proof of Permutation Formula) To arrange r items from n distinct items:

- n options for the first position.
- $n - 1$ for the second.
- Continue until $n - r + 1$ options for the r-th position.

Thus,

$$\begin{aligned}
P(n, r) &= n \times (n-1) \times \cdots \times (n-r+1) \\
&= \frac{n!}{(n-r)!}.
\end{aligned}$$

□

Remark

- If $r = n$, then $P(n, n) = n!$.
- If $r = 0$, then $P(n, 0) = 1$.

Example (Seating 5 People) How many ways can 5 people be seated in a row?

$$P(5, 5) = \frac{5!}{(5-5)!} = 5!.$$

Thus, there are $5! = 120$ ways.

Definition 12 (Combination) An arrangement of r objects from n objects without considering the order.

$$C(n,r) = \frac{n!}{r! \times (n-r)!}.$$

PROOF (Combination Formula) Number of ways to choose r items from n distinct items without considering the order. First find the number of ways to arrange r items from n distinct items, then divide by the number of ways to arrange the r items. Therefore, the total

$$\begin{aligned} C(n,r) &= \frac{P(n,r)}{P(r,r)} \\ &= \frac{n!}{r! \times (n-r)!} \times \frac{r!}{r!} \\ &= \frac{n!}{r! \times (n-r)!}.. \end{aligned}$$

□

Remark

- $C(n,r) = C(n,n-r)$.
- $C(n,0) = 1$.
- $C(n,1) = n$.
- $C(n,n) = 1$.

Example An 8-bit codeword is selected at random. What is the probability that it contains at least 3 zero bits?

$$\begin{aligned} T &= \sum_{i=3}^{8} \binom{8}{i} \\ &= 2^8 - \sum_{i=0}^{3} \binom{8}{i} \\ &= 163 \\ P &= \frac{163}{2^8} = 0.6367 \end{aligned}$$

.

Subsection 2.3

Binomial Theorem

Definition 13 (Binomial Theorem)

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k.$$

Example Expand $(x+y)^3$.

$$\begin{aligned} (x+y)^3 &= \binom{3}{0} x^3 y^0 + \binom{3}{1} x^2 y^1 + \binom{3}{2} x^1 y^2 + \binom{3}{3} x^0 y^3 \\ &= x^3 + 3x^2 y + 3xy^2 + y^3. \end{aligned}$$

Section 3

Conditional Probability

Definition 14 (Conditional Probability) The probability of an event B occurring given that A has occurred. It is denoted as $P(B|A)$.

$$P(B|A) = \frac{P(A \cap B)}{P(A)}; \quad P(A) > 0. \tag{3.1}$$

similarly,

$$P(A|B) = \frac{P(A \cap B)}{P(B)}; \quad P(B) > 0.$$

In a Venn Diagram, conditional probability is equivalent to changing the sample space to B and calculating the probability of A in this new space. For example, in Figure 4, $P(A|B) = \frac{P(A \cap B)}{P(B)}$. Imagine wrapping the space around B and considering only the outcomes in this new space.

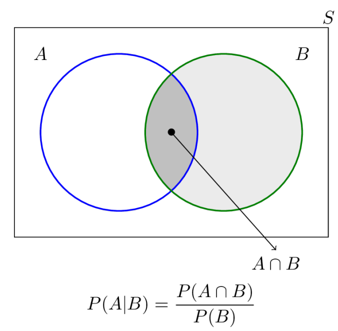

Figure 4. Conditional Probability in a Venn Diagram.

Definition 15 (Independence of Events) Two events A and B are independent if:

$$P(B|A) = P(B) \quad \text{or} \quad P(A|B) = P(A).$$

Equivalently:

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

Independence is not the same as disjointness. Disjoint events are dependent, but independent events are not disjoint. Not related to mutual exclusivity.

Theorem 1 (General Multiplicative Rule) For events $A_1, A_2, \ldots, A_k$:

$$P(A_1 \cap A_2 \cap \cdots \cap A_k) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \cdots P(A_k|A_1 \cap \cdots \cap A_{k-1}). \tag{3.2}$$

$$P(\bigcap_{i=1}^{k} A_i) = \prod_{i=1}^{k} P(A_i|\bigcap_{j=1}^{i-1} A_j). \tag{3.3}$$

If the events are independent:

$$P(A_1 \cap A_2 \cap \cdots \cap A_k) = P(A_1)P(A_2) \cdots P(A_k). \tag{3.4}$$

Proof (General Multiplicative Rule) To prove the formula for $P(A_1 \cap A_2 \cap \cdots \cap A_k)$, we proceed by using the definition of conditional probability iteratively.

$$P(A_1 \cap B_1) = P(B_1)P(A_1|B_1)$$

Let $B_1 = A_2 \cap B_2$,

$$P(A_1 \cap A_2 \cap B_2) = P(A_2 \cap B_2)P(A_1|A_2 \cap B_2)$$

$$= P(B_2)P(A_2|B_2)P(A_1|A_2 \cap B_2)$$

Since intersection is associative, we can reorder the terms.

$$P(B_2 \cap A_2 \cap A_1) = P(B_2)P(A_2|B_2)P(A_1|A_2 \cap B_2)$$

Change of variables,

$$P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2)$$

Repeat k-times,

$$P(A_1 \cap A_2 \cap \cdots \cap A_k) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \cdots P(A_k|A_1 \cap \cdots \cap A_{k-1}).$$

$\square$

Theorem 2 (Total Probability) If events $B_1, B_2, \ldots, B_k$ partition the sample space S:

$$P(A) = \sum_{i=1}^{k} P(B_i \cap A) = \sum_{i=1}^{k} P(B_i)P(A|B_i). \tag{3.5}$$

Total probability is a way to calculate the probability of an event A by considering all possible ways it can occur. In Figure 5, the probability of A is the sum of the probabilities of A given each event B_i times the probability of B_i. We usually trace the path of A through the events B_i to calculate the conditional probability using a tree diagram (Figure 6) to reach the total probability.

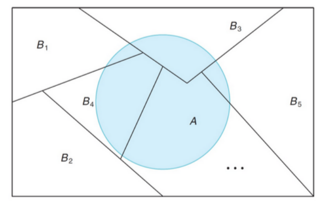

Figure 5. Total Probability in a Venn Diagram.

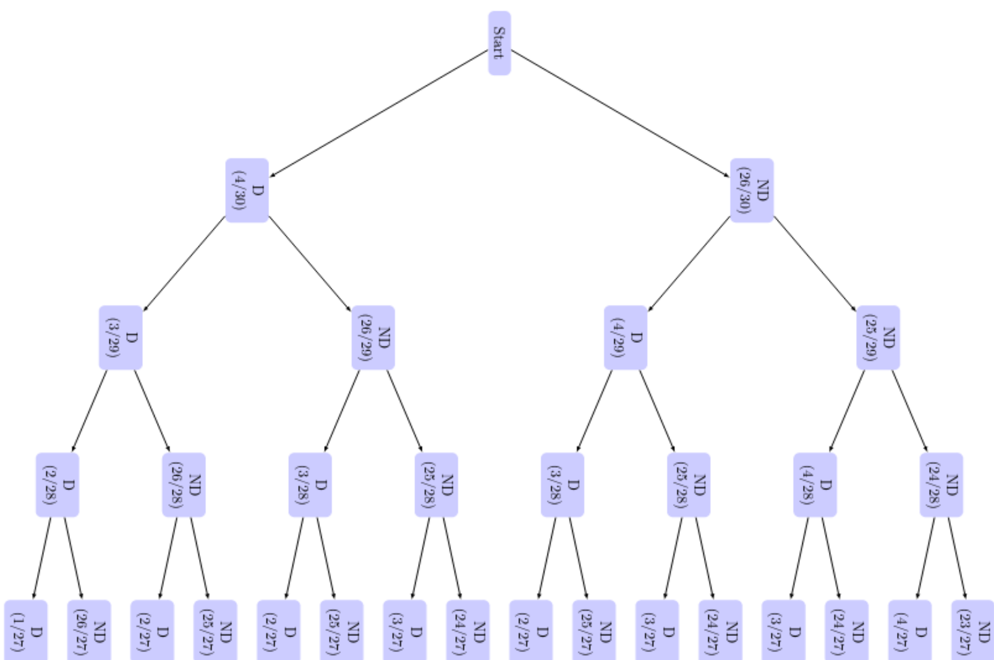

Figure 6. Tree Diagram for Conditional Probability.

Theorem 3 (Bayes' Rule) If events $B_1, B_2, \ldots, B_k$ partition the sample space S:

$$P(B_r|A) = \frac{P(A|B_r)P(B_r)}{\sum_{i=1}^{k} P(A|B_i)P(B_i)}. \tag{3.6}$$

Proof (Bayes' Rule)

By the definition of conditional probability 14,

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

By the multiplication rule 1,

$$= \frac{P(A|B_r)P(B_r)}{\sum_{i=1}^{k} P(A|B_i)P(B_i)}.$$

□

PART II

Random Variables

Section 4

Discrete Random Variables

Subsection 4.1

Introduction

A *discrete random variable (RV)* is a type of variable that can take on a finite or countably infinite set of values. It provides a numerical summary of outcomes from a random experiment.

Example (Coin Toss) Consider tossing a coin twice. Define the random variable X as the number of heads observed. The range of X is $\{0, 1, 2\}$.

Example (Switchboard Calls) Let X denote the inter-arrival time between calls and Y the number of calls received in a day at a switchboard. The range of X is the set of non-negative real numbers, while the range of Y is the set of non-negative integers.

Subsection 4.2

Probability Mass Function (PMF)

The *probability mass function (PMF)* of a discrete RV X describes the probability of each possible value:

$$f(x) = P(X = x),$$

subject to the following properties:

1. $f(x) \geq 0$ for all x.
2. $\sum_x f(x) = 1$.

Example (Defective Computers) A shipment of 8 microcomputers includes 3 defective units. If a random selection of 2 computers is made, determine the probability distribution of the number of defective computers selected.

Subsection 4.3

Cumulative Distribution Function (CDF)

The *cumulative distribution function (CDF)* of a discrete RV X is given by:

$$F(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i),$$

with the following properties:

1. $F(x)$ is non-decreasing.
2. $F(x) \in [0, 1]$ for all x.

Example (Coin Toss Until a Tail) A coin is tossed until a tail appears or three attempts are made. Determine the PMF and CDF for the number of tosses required, and sketch their graphs.

Subsection 4.4

Expected Value (Mean)

The *expected value* or *mean* of a discrete RV X is defined as:

$$\mu = E[X] = \sum_x x f(x).$$

Example (Expected Number of Chemists) A committee of size 2 is randomly selected from a group of 4 chemists and 3 biologists. Compute the expected number of chemists on the committee.

Subsection 4.5

Variance

The *variance* of a discrete RV X measures its spread around the mean:

$$\sigma^2 = \mathrm{Var}(X) = E[(X - \mu)^2].$$

Example (Revenue Comparison) Two product designs are compared based on revenue:

- Design A has a fixed revenue of \$3 million.
- Design B has a 30% chance of yielding \$7 million and a 70% chance of yielding \$2 million.

Calculate the mean and standard deviation for each design.

Subsection 4.6

Expected Value of a Function

For a function $g(X)$ of a discrete RV X, the expected value is:

$$E[g(X)] = \sum_x g(x) f(x).$$

Example | (Linear Transformation) Find $E[aX + b]$ for constants a and b.

Stochastic Processes

PART

III